Resumo:

As esponjas marítimas são metazoários básicos distribuídos por vários ambientes aquáticos. São conhecidas por possuírem comunidades microbianas densas e complexas de diferentes linhagens. Apesar do progresso obtido no estudo do consórcio microbiano, existe pouca informação disponível sobre os mecanismos envolvidos na relação simbiótica entre as bactérias e as esponjas-hospedeiro. De forma a perceber as capacidades funcionais de um organismo simbiótico, foi analisado o genoma de um organismo simbiótico, *Pseudovibrio* sp. Poly-S9, isolada da esponja marinha *P. penicillus.* As análises genómicas efetuadas revelaram um genoma excecionalmente grande de 6.6 Mbp, relativamente a outros dois organismos do mesmo género já relatados, FO1-BEG e JE062, isoladas de um coral e da esponja marinha *Mycale laxíssima*, respetivamente. Informação genómica adicional revelou a capacidade da bactéria se adaptar a em resposta a várias condições do seu hospedeiro e os mecanismos moleculares que poderão afetar a interação proto-eucarióta. O simbionte possui um reportório de genes que codificam proteinas semelhantes às dos eucariotas, entre as quais ANKs, TPRs e domínios Sel1 que medeiam as interações proteína-proteina, interferindo nas funções celulares do hospedeiro. Foi também encontrado neste genoma um numero elevado de agentes de transferência de genes (GTAs) indicando a presença de estratégias hospedeiro-adaptativas e eventos frequentes da transferência horizontal de genes (HGT). Este estudo pretende fornecer informação para a compreensão dos mecanismos moleculares envolvidos em simbioses proto-eucariotas.

Palavras-chave: Esponja do mar; Simbiose; Secreção; NRPS-PKS; ELP

Summary:

Sponges are one of the ancient basal metazoans distributed in various aquatic environments. They are known to harbour dense and complex microbial communities of different lineages. Sponge-associated microbes live symbiotic association with the sponge hosts benefiting each other through the integration of metabolites. Elucidating the mechanisms involved in establishing a successful microbial association with a sponge partner through the genome analyses has been the focus of this study. The genome analysis of *Pseudovibrio* sp. POLY-S9 strain isolated from the intertidal marine sponge *Polymastia penicillus* revealed a large genome size of 6.6 Mbp compared to the previously reported genomes of *Pseudovibrio*. Detailed genome mining detected a genetic repertoire containing, non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS), among others such as bacteriocin or terpenes suggesting the capability of POLY-S9 to produce a wide variety of metabolites. Genomic insight further confirmed the presence of eukaryotic-like proteins such as the ankyrin-repeat proteins (ARPs), tetratrico peptide repeats (TPRs) and Sel1 repeats that involve in altering the host behavior through protein-protein interactions. The genome of Pseudovibrio POLY-S9 harbored higher proportion of mobile DNA elements (~5%) and gene transfer agents (GTA), suggesting the massive genome expansion and permissive lifestyle through horizontal gene transfer (HGT). The enhanced horizontal gene transfer (HGT) pathways among the strain POLY-S9 suggests a key role in distributing the core functions responsible for the symbioses during their co-evolutionary association with their hosts and the ability to harbour a wide range of eukaryotic organisms. In conclusion, the genome of POLY-S9 exhibited an increase in size, number of mobile-DNA, multiple metabolite gene clusters and secretion systems, likely to influence the genome diversification and the adaptability.

Keywords: Sponge; Symbioses; Secretion, NRPS-PKS; ELP

Index

Figures and Tables Index

Abbreviation List

NRPS - Non-ribosomal peptide synthetase (NRPS

PKS - Polyketide synthase

APR – Ankyrin-repeat proteins

TPR – Tetratrico peptide repeats

GTA – Gene transfer agents

HGT – Horizontal gene transfer

ELP – Eukaryotic-like proteins

CDS – Coding sequences

rRNA – Ribosomal RNA

tRNA – Transfer RNA

COG – Cluster of orthologous groups

T3SS – Type III secretion system

T4SS –Type IV secretion system

T6SS – Type VI secretion system

AHL - N-acyl homeserine lactones

Introduction:

Symbioses: An overview

Symbiosis is considered as a long-term association between two organisms of different species where they live together for their mutual benefit. (Bennett, 1887). In nature, some symbiotic relations are obligatory because the survival of the host and the partner depends on each other. Whereas in the facultative relationship the beneficiary effect of the host and the partner happens in a given time or conditions. It is known that in many cases, more than one symbiont can be found in single host. The main question is "*why did these two different organisms live together?* Or *how did this association started?"* The answer is not so simple because each pair of host-symbiont has its own reasons for existence.

Bennett first defined symbiosis when talking about lichens, a most common symbiotic association between the fungi and the photosynthetic bacteria (Oxford, 2005). Symbiotic relationships are evident in both terrestrial and aquatic ecosystems, representing fresh and marine water. Symbioses can be extremely helpful in enhancing the host's fitness in order to survive in a specific habitat. Different environmental advantages may be conferred to host depending on the symbiont's characteristics. During the course of evolution these associations resulted in the development of more complex and essential compartments likely mitochondria and chloroplast.

In the course of time the symbiont loses its genes and lack its ability to live in isolation. So, how is that we may still observe symbioses that started over 2 billion years ago? The answer to this question is not very clear, but evidences point to the ability of the symbiont to also insert its genes in germ cells and hence reproduce and then also "infecting" its descendants. This mode of transfer is termed vertical transmission. Horizontal transmission occurs, for example, when a symbiont reproduces itself in a different rate than its host and its descendants are secreted to the environment to find a host partner themselves, like *zooxathellae* symbiotic association with corals.

Symbioses in marine habitat: Cnidaria

In marine environment symbioses has been well documented in Sponges, Cnidarians and Ascidians (Fig.1)

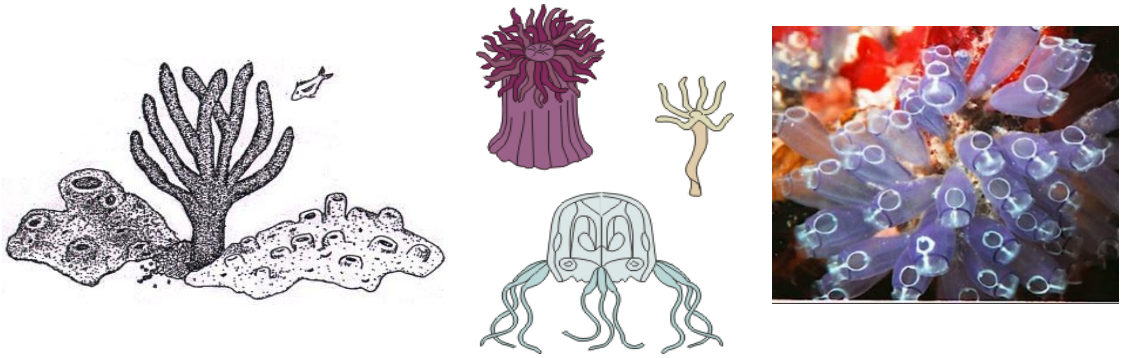


Figure 1- Three most widely studied marine invertebrates for the symbiotic association, sponges (A), Cnidaria (B) and Ascidians (C).

The symbiotic association between the photosynthetic dinoflagellates (*Symbiodinium*) and Cnidaria is very common in coral reefs. Cnidaria-dinofalgellate symbioses centered on nutrient exchange where the dinoflagellates supply high amount of organic matter produced via photosynthesis in return of inorganic nutrients such as nitrogen and carbon, a high light environment and refuge from predators (Muller-Parker and D'Elia, 1997).

Symbiosis in ascidians

Ascidians are found throughout all oceans. (Schmitt et al., 2012; Shenkar and Swalla, 2011). Symbiosis is highly evident in the production of secondary metabolites with defensive role among Ascidians (Joullié et al., 2003; Paul et al., 1990). It has been reported that these secondary metabolites with defensive role are produced not only by the animal host but also by its symbiotic bacteria(Donia et al., 2006, 2008, 2011a, 2011b; Kwan et al., 2012; Rath et al., 2011; Schmidt and Donia, 2009). Defense enhancing metabolites range from toxic cyclic peptides, cyanobactins (Donia et al., 2008) in Family *Didemnidae* as well as UV protectant by mycosporine (Donia et al., 2011c; Hirose and Fukuda, 2006). Some obligatory intracellular Alphaproteobacteria produce highly toxic polyketides (Kwan et al., 2012). This is a long-term association dating from at least 6 million to at most 35 million years(Kwan and Schmidt, 2013; Kwan et al., 2012). Polyketides are also found in an Antarctic didemnid Ascidian (Riesenfeld et al., 2008). Rath et al in 2011 also reported that a gammaproteobacteria produces potent ecteinascidians within the non-didemnid ascidian *Ecteinascidins turbinate*.

Sponge microbe association

Sponges (Phylum Porifera) are widely distributed and are abundant in most marine ecosystems(Diaz and Rützler, 2001), and are well known to harbour a wide range of microbial communities (Table 1). Sponge belonging to the class *Demospongiae* are widely studied due to its prominent distribution ranging from fresh water to marine. Sponges are sessile filter feeders with a simple body plan with the ability to channel water through the body cavity and osculum (Fig. 4). These lower invertebrates are known for comprising 50% of the body mass with the different microbial community(Taylor et al., 2007). The simple body plan and the larger surface quotient to volume made the symbionts to escape the host immune system. Sponges are distribute in 4 classes: Demospongiae, Hexactinellida, Calcarea and Homoscleromorpha. Further lower taxonomic division allowed the classification into 25 orders, 128 families and 680 genera. Class Demospongiae is the most widespread with approximately 83% of the 8553 formally recognised sponges (Van Soest et al., 2012)

Demospongiae usually forms Leuconoids in structure. Skeleton constitute spicules made of sponging, mineral silica or both.

Figure 2- *Hymeniacidon perlevis*, a sponge from Demospongiae class

Sponges under class Calcarea form Asconoid, syconoid or leuconoid structure (Fig. 3). They are strictly marine sponges and their spicules are made from calcite or aragonite. They present themselves as small in height and drab in color. Their body is usually vase shaped and the skeleton is either a mesh or honeycomb structure.

Figure 3- A sponge from calcarea class- Copyright Valentino Valicelli

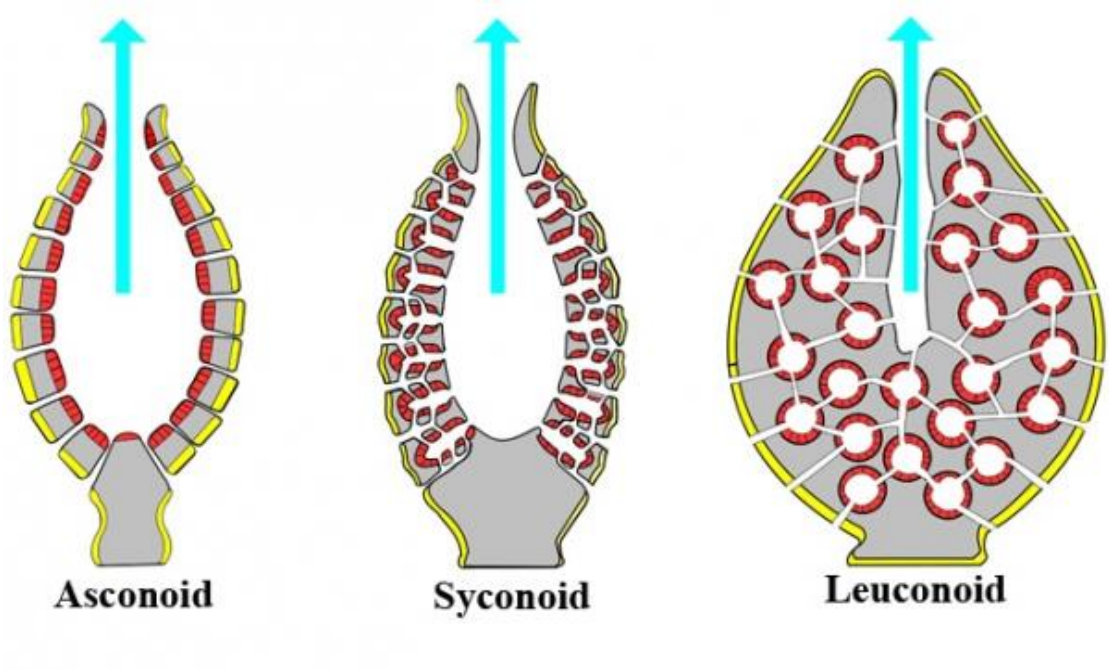


Figure 4 - Different body forms of sponges.

Hexactinellida sponges are known as glass sponges due to their colour and shape. The skeleton consisted of 4 or 6 siliceous spicules. Most of the sponges of the class Hexactinellida are found in deep ocean depths. They are scattered around the world but they are mostly present in Antarctic and north Pacific oceans.(Fondo and Martens, 1998). Their body has symmetry and a large cavity formed by the skeleton. Unlike other sponges, their cytoplasm is not divided by walls but instead it forms a syncytium with many nuclei (Leys et al., 2007). Some species may fuse together to create bioherms.

Figure 5 - *Euplectella aspergillum,* a Hexactinellida sponge

Homoscleromorpha is the most recent addition to the sponges classes (Gazave et al., 2010, 2012)(Gazave et al, 2010; 2012). These sponges are massive or encrusting bodies and have a very simple structure composed of very small spicules.

Figure 6 - *Oscarella lobularis,* a Homoscleromorpha sponge

Most of the sponge-associated microbial studies directed either towards estimating the diversity or the biogeography pattern of the microbes. But two studies (Taylor et al., 2007; Vogel, 2008) have pointed out the significance of the density and diversity of resident microbes in sponges' function and its utility to pharmacologists and biotechnologists. Initially, it was thought that biomolecules produced were sponge metabolites, but further research showed that the sponge-associated microbes were responsible for the production of those compounds There is little direct evidence that confirms role of symbionts to the well-being or survival of the sponge, except in the case of *cyanobacteria* (Wilkinson and Fay, 1979) in which sponge's health declined with the loss of cyanobiont (Thacker, 2005). For instance, *Cliona varians* when exposed to

bleaching events are able to acquire symbionts from the environment (Hill and Wilcox, 1998). However, the different symbionts with different metabolic pathway show circumstantial evidences of improving sponge health like, ammonium-oxidazing archaea (Steger et al, 2008), nitrite-oxidizing phylum Nitrospira (Hentschel et al., 2001) sulphate reducing bacteria (Hoffmann et al., 2005), anaerobic phototrophs (Imhoff and Trüper, 1976) and even vertical transfer from parents to larval offspring (Steger et al., 2008). However, the interactions at the molecular level are not yet clear. The role of microbes in sponge biology and ecology is considered a Holy Grail in microbial ecology as there is no indication of "who eats what and when?" (Neufeld et al., 2007). Understanding the language spoken between host and symbiont should reveal how symbiosis occurs and what compels the association to persist.

Table 1. Sponges and their natural products showing various bioactivity

Sponge	Bioactive metabolites	Biological activity
Acanthella sp.	Kalihinol-A	Antibiotic
Agelas dispar	Aminozooanemonin	Antibacterial
Agelas dispar	Pyridinebetaine A	Antibacterial
Agelas mauritiana	Agelasimine	Cytotoxic
Agelas mauritiana	Sceptrin	Antimicrobial
Agelas nakamurai	Ageliferine	Antibacterial
Agelas nakamurai	Debromosceptrin	Antibacterial
Agelas nakamurai	Nakamuric acid	Antibacterial
Agelas novaecaledoniae	Ageliferine	Somatostatin/VIP inhibitor
Agelas novaecaledoniae	Sceptrin	Somatostatin/VIP inhibitor
Agelas novaecaledoniae	Xestospongine B	Somatostatin/VIP inhibitor
Agelas sp.	Agelasine	Antileukemic
Agelas sp.	Agelasine F	Antituberculosis
Agelas sp.	Agelasine I	Antimicrobial
Amphimedon sp.	Pyrinodemin	Cytotoxic
Aplysina aerophoba	Aeroplysinin I	Cytotoxic
Batzella sp.	Discorhabdin	Cytotoxic, enzyme inhibitor
Batzella sp.	Secobatzelline	Phosphatase inhibitor
Crella sp.	Crellastatins	Cytotoxic
Corticium sp.	Meridine	Antifungal
Cymbastela sp.	Agelastatins C, D	Insecticidal
Discodermia calyx	Calyculin A	Antitumor
Discodermia kiiensis	Discodermin A	Antimicrobial
Discodermia sp.	Discobahamins	Antifungal
Disidea avara	Avarol	Cytotoxic
Druinella purpurea	Psammaplysin C	Cytotoxic
Dysidea sp.	Furodysinin	Antiparasitic
Echinoclathria sp.	Echinoclathrines	Immunosuppressive
Erylus lendenfeldi	Eryloside A	Antitumor, antifungal
Fascaplysinopsis reticulata	β-Carbolium salt	Antiparasitic
Halichondria okadai	Halichondrin B	Antitumor
Haliclona osiris	Osirisynes	Na^+/K^+-ATPase inhibitor

Haliclona sp.	Manzamine A	Antitumor
Haliclona tulearensis	Halitulin	Cytotoxic
Hamacantha sp.	Hamacanthin	Antifungal
Hyrtios erecta	Heteronemin	Antiparasitic
Ianthella basta	Bastadin	Antimicrobial
Ianthella sp.	34-sulfatobastadin 13	Endothelin A receptor inhibitor
Ircinia sp.	Haterumalides	Cytotoxic
Jaspis johnstoni	Jasplakinolide	Cytotoxic
Jaspis johnstoni	Jasplakinolide	Insecticidal
Jaspis johnstoni	Toyocamycin	Cytotoxic
Jaspis johnstoni	Tubercidin	Cytotoxic
Jaspis sp.	Bengamides	Antitumor
Jaspis sp.	Bengazoles	Antiparasitic
Jaspis sp.	Cyclodepsipeptide	Antifungal
Jaspis sp.	Jaspisamides	Cytotoxic
Jaspis sp.	Psammaplin	Antibacterial
Jaspis splendans	Jaspamide	Antitumor
Jaspis wondoensis	Wondosterols	Antimicrobial
Latrunculia magnifica	Latrunculin A	Neurotoxin
Leucetta cf. chagosensis	Isonaamidine D	Antifungal
Neosiphonia superstes	Sphinxolides	Cytotoxic
Notodoris citrina-Leucetta chagosensis	Naamidines & naamines	Antiparasitic

Table 1. Continued

Sponge	Bioactive metabolites	Biological activity
Pachastrissa sp.	Bengamides	Antifungal
Pachastrissa sp.	Bengazoles	Antifungal
Pandaros acanthifolium	Acanthifolicin	Antitumor
Petrosia sp.	Petrocortynes	Cytotoxic, enzyme inhibitor
Petrosia sp.	Petrotetrayndiols	Cytotoxic
Petrosia sp.	Petrosiacetylenes	Na^+/K^+-ATPase inhibitor
Plakinastrella sp.	Elenic acid	Topoisomerase II ihibitor
Poecillastra wondoensis	Wondosterols	Antimicrobial
Psammaplysilla crassa	Purealin	Antiparasitic
Psammaplysilla purpurea	Aeroplysinin I	Antiparasitic
Psammaplysilla purpurea	Bastadin	Antimicrobial
Psammaplysilla purpurea	Purealidin A	Cytotoxic
Reidispongia coerulea	Reidispongiolide	Cytotoxic
Reniera cratera	Dorimidazole A	Antiparasitic
Rhaphisia lacazei	Topsentins	Antiproliferative
Spongia sp.	Spongianolide	Cytotoxic
Spongionella gracilis	Gracilin B	Cytotoxic
Stronglyophora hartmani	Puupehenone	Cytotoxic
Stylinos n. sp.	Mycalamides	Cytotoxic
Suberea creba	Aeroplysinin I	Antibacterial
Suberea creba	Dibromoverongiaquinol	Antibacterial
Tedania digitata	1-methylisoguanosine	Cardiovascular effector
Tedania ignis	Tedanolide	Cytotoxic
Tethya crypta	Spongouridine, Spongothymidine	Antiviral, antitumor
Theonella sp.	Koshikamide	Cytotoxic
Theonella swinhoei	Swinholide	Antifungal
Theonella swinhoei	Theopederins	Antifungal, cytotoxic
Verongia aerophoba	Aeroplysinin I	Antibacterial
Verongia aerophoba	Dienone	Cytotoxic
Verongia spengelii	Aplysinopsin	Cytotoxic
Xestospongia sp.	Xestospongine B	Somatostatin/VIP inhibitor
Xestospongia sp.	Ageliferine	Somatostatin/VIP inhibitor
Xestospongia sp.	Sceptrin	Somatostatin/VIP inhibitor
Xestospongia sp.	Xestoaminol A	Antiparasitic
Zyzzya fuliginosa	Veiutamine	Cytotoxic

So far, the microbial ecology could answer the diversity of uncultivated microbes live in association with the invertebrate hosts. Whole genome sequencing of the microbes isolated from the sponges help to understand the molecular mechanisms involved in the host-microbe association.

Natural Marine Products research and their importance

Over the last two decades, the discovery of marine natural products (NMNP) has increased gradually from 20 compounds to 200 compounds per year (Hu et al., 2011) (Hu et al, 2011). By the year 2010, it was estimated that more than 15 000 new marine natural products had been discovered (Blunt et al., 2007; Faulkner, 2000, 2001).

As stated before, marine sponges (Porifera) are recognized as the richest sources of NMNP, contributing with about 30% of the all marine natural products so far. It is mandatory to refer that from 4851 compounds supplied from sponges, 1499 were isolated in recent years between 2008 and 2012 (Blunt et al, 2010-2014). This indicates that the sponges were the most prolific supplier of NMNP in the last decade. (Laport et al., 2009)

With the rising number of NMNP, a broad range of bioactivities ranging from anticancer, antiviral, antibacterial, antifungal, antiprotozoal, anthelmintic, anti-inflammatory, immunosuppressive, neurosuppresive, neuroprotective and antifouling were reported (Blunt et al., 2006). With the evolution of pathogens and growing resistance to current drugs, marine sponges have provided novel leads against diseases (Blunt et al., 2006; Laport et al., 2009; Sagar et al., 2010). If the current rate is maintained, most therapies for human diseases may be based on NMNPs. (Koopmans and Wijffels, 2008; Molinski et al., 2009; Paul et al., 2011)

In their natural habitats, sponges have to have the ability to defend themselves and to compete for nutrients. Most marine benthic environments harbor sponges but also a very broad range of other marine organisms. Sponges are simple organisms that lack locomotion and therefore, they have to have another mechanism for environmental advantage over their counterparts and avoid predator attacks, microbial infections, biofouling and the growth of other sessile organisms (Paul and Puglisi, 2004,Paul et al., 2006). In order to unravel the fascinating mechanisms of sponge-microbe interactions and due to the sheer number of promising novel pharmaceutical compounds used in the treatment of human diseases, sponges are very attractive for researchers (Attaway and Zaborsky, 1993).

Sponge mesohyl is often inhabited by microbes and some natural products retrieved from marine microbes resemble the ones isolated from sponges (Mcclintock et al, 2010). Sponges harbour microbial communities that exceed their surroundings by 2 or 3 magnitude orders. The sponge mesohyl, is considered a unique niche where majority of the bacterial community (30% to 60%) resides (Selvin et al, 2009).

Bacteria-specific traits are provided to the host, being the most notable autotrophy, nitrogen fixation and nitrification (Wilkinson and Fay, 1979). Other features such as metabolic waste processing, skeleton stabilization and UV radiation protection are also reported. (Diaz and Rützler, 2001; Lee et al., 2011; Thoms et al., 2003).The intensive interaction with the environment has given a unique biochemistry to sponges besides prompting them to survive or evolve (Proksch, 1994). Some chemical prevent the colonization of some possible damaging organisms (Mahon et al., 2003; Paul et al., 2001, 2001) (Paul et al, 2006; Mahon et al, 2003; Paul et al, 2001) and defense mechanism against viral, parasitic and fungal diseases (Unson et al., 1994).

Pseudovibrio sp.:

Pseudovibrio sp. was initially isolated from coastal seawater. It was considered to be a marine, heterotrophic, and facultative anaerobic bacterium capable of denitrification and fermentation. Shieh et al (2004) named it *Pseudovibrio denitrificans*. Strains *P. ascidiaceicola* (Fukunaga et al., 2006)*, P. japonicus* (Hosoya and Yokota, 2007) and *P. axinellae* (O'Halloran et al., 2011) were also described. However, some *Pseudovibrio* related bacteria were detected around the world in various studies while analyzing the 16S rRNA (Agogué et al., 2005; Alex and Antunes, 2015; Chiu et al., 2012; Enticknap et al., 2006; Flemer et al., 2012; Hentschel et al., 2001; Kennedy et al., 2009; Koren and Rosenberg, 2006; Lafi et al., 2009; O'Halloran et al., 2011; Olsén et al., 2002; Penesyan et al., 2011; Riesenfeld et al., 2008; Santos et al., 2010; Sertan-de Guzman et al., 2007; Thakur et al., 2003; Thiel and Imhoff, 2003; Thoms et al., 2003). With the exception of a few strains, *P. japonicus* and *P. ascidiaceicola* were reported to be an epibiont of a red algae (Penesyan et al., 2011) and an isolate from costal oligotropic seawater (Agogué et al., 2005). All other strains were reported from marine invertabrates, mainly sponges but also coral and tunicates. *Pseudovibrio* is known to dominate the bacterial culture isolates, like in the cases of sponge *Rhopaloeides odorabile*, *Suberites domuncula*, *Clathrina clathris*, and others (Webster et al., 2012) and are also considered the most dominant in larvae of sponge, *Mycale laxissima*, indicating the vertical transmission of these bacteria (Enticknap et al., 2006)

The widespread presence of *Pseudovibrio* sp. in various habitats and its ability to thrive as free-living bacteria, as an algal epiphyte and as a symbiont with other invertebrate hosts prompted to investigate the genomics nature of the isolated bacterium. In this thesis we performed the whole genome sequencing of the *Pseudovibrio* sp. isolated from the intertidal sponge *Polymastia penicillus* (class Demospongiae) sampled from the Atlantic coast of Portugal to investigate the evolutionary genomics mechanism favoring the adaptation of this bacterium to thrive within the sponge host. The draft genome of the *Pseudovibrio* sp. POLY-S9 revealed the presence of a genomic repertoire (genes for protein translocation, attachment, and invasion; secretion system apparatus and its effector molecules) responsible to interfere with intrinsic host cell function and bacterial survival. Genome analyses also revealed the presence of various secondary metabolic gene clusters. The large number of mobile/gene transfer agents suggests the genome plasticity of the bacterium and mechanism of gene procurement through horizontal gene transfer.

In this thesis we investigated the molecular mechanisms involved in the proto-eukaryote symbioses. We sequenced the whole genome of *Pseudovibrio* sp. POLY-S9 isolated from marine sponge, *Polymastia penicillus* and comparative genomics analyzes was performed with previously sequenced genome of the genus *Pseudovibrio* isolated from different environments.

Methods:

Genome sequencing/retrieval, assembly and annotation

The genomes of three different strains of the *Pseudovibrio* (POLY-S9, FO-BEG1 and E062,) were analyzed in this study. Among them, *Pseudovibrio* sp. POLY-S9 was isolated previously during my dissertation (Silva, 2012) from the intertidal marine sponge *Polymastia penicillus* sampled from the Atlantic coast of Portugal (Alex et al, 2013). Genome sequencing and assembly of *Pseudovibrio* sp. POLY-S9 were performed by Alex et al (unpublished).

The genome of *Pseudovibrio* sp. POLY-S9, was annotated using the Prokka v1.10 (**Seemann 2014**) annotation pipeline. Prokka depends on many external feature prediction tools for genome annotation. Coding sequence (CDS) predication was carried out by Prodigal v2.60 (**Hyatt et al. 2010**). We compiled a core database for assigning the function to the predicted CDS using a set of trustworthy curated dataset comprising annotated bacterial proteins obtained from the current release of UniProt Knowledgebase, UniProtKB Release 2014_10 (**Magrane & Consortium 2011**). Ribosomal RNA (rRNA), transfer RNA (tRNA) genes and non-coding RNA were

predicted using RNAmmer v1.2 (**Lagesen et al. 2007**), ARAGORN v1.2.36 (**Laslett & Canback 2004**), and Infernal v1.1 (**Kolbe & Eddy 2011**), respectively. SignalP v4.1 was used for the prediction of signal peptides (**Petersen et al. 2011**). The genomes of the genus *Pseudovibrio* (FO-BEG and JE062) used for comparative purposes in this study were downloaded from NCBI ftp (ftp://ftp.ncbi.nih.gov/genomes/). The sequencing, assembly and annotation of FO-BEG and JE062 are reported in the previous study (**Bondarev V et al. 2013**)

16S rRNA phylogeny of the genus *Pseudovibrio*

We used a dataset comprising the 16S rRNA gene sequences of the *Pseudovibrio* sp. POLY-S9 sequenced in this study and its closest BLAST hit, and the 16S sequences of other *Pseudovibrio* sp. used in the previous study (**Bondarev et al. 2013**) representing the strains isolated from different sources. 16S rRNA sequences of Sequences were aligned with clustalW2 (**Larkin et al. 2007**) and filtered out the alignments using GBlocks (**Castresana 2000; Talavera & Castresana 2007**). Sequences were manually edited to remove the gaps and other ambiguous sequences using Jalview V.2 (**Waterhouse 2009**). Maximum-likelihood phylogenetic tree was constructed in PhyML (**Guindon & Gascuel 2003**) using the best fit evolutionary model, GTR+G+I adopted from Akaike Information Criterion with correction (AICc) implemented in MrAIC ver. 1.4.4 (**Nylander 2004**).

Secondary Metabolite gene clusters identification:

NRPS and PKS search was performed through antibiotic and secondary Metabolites Analysis SHell (antiSMASH) V. 3.03 (Webber et al, 2015)

Function classification with Cluster of Orthologs Groups (COG) analysis:

Cluster of Orthologs Groups (COG) analysis was performed using the webMGA (http://weizhong-lab.ucsd.edu/metagenomic-analysis) database with the *e*-value of $1x10^{-5}$. Results were visualized via Microsoft® Office™ Excel™ 2013. Phage DNA sequences within the bacterial genome were identified using PHAST (Zhou et al., 2011).

Circular genome maps:

Circular genome maps were designed using BLAST Ring Image Generator (BRIG v.0.95) (Alikhan et al, 2011). GC content and GC skew were added with the same software. Features of interest were added using the “Custom feature option” of the same software.

Results and Discussion

General genome features:

The isolated *Pseudovibrio* sp. POLY-S9 from the intertidal marine sponge *P. penicillus* showed 100% SSU rRNA similarity to the *P. ascidiaceicola* strain NBRC 100514 (AB681198), which has been previously isolated from the ascidian *Polycitor proliferus* (Fukunaga et al., 2006). Phylogenetic analyses further revealed the clustering of POLY-S9 strain with *P. ascidiaceicola* isolated from tunicates (AB681198, AB175663, AB210280) suggesting the ability of the genus *Pseudovibrio* to infect a wide range of marine hosts.

The genome analyses of the strain *Pseudovibrio* sp. POLY-S9 isolated from the intertidal marine sponge *P. penicillus* revealed the chromosomal genome size of 6.6 Mbp with a G + C content of 51.26% (table 2 and figure 9) consistent with previously described values for the *P. ascidiaceicola* strain (Fukunaga et al., 2006). The genome assembly of the *Pseudovibrio* sp. POLY-S9 resulted in 271 contigs and the annotation with Prokka revealed the presence of 6171 coding sequences (CDS), greater than the estimated CDS from the genomes of *Pseudovibrio* sp. FO-BEG1 and *Pseudovibrio* sp. JE062 (table 2).

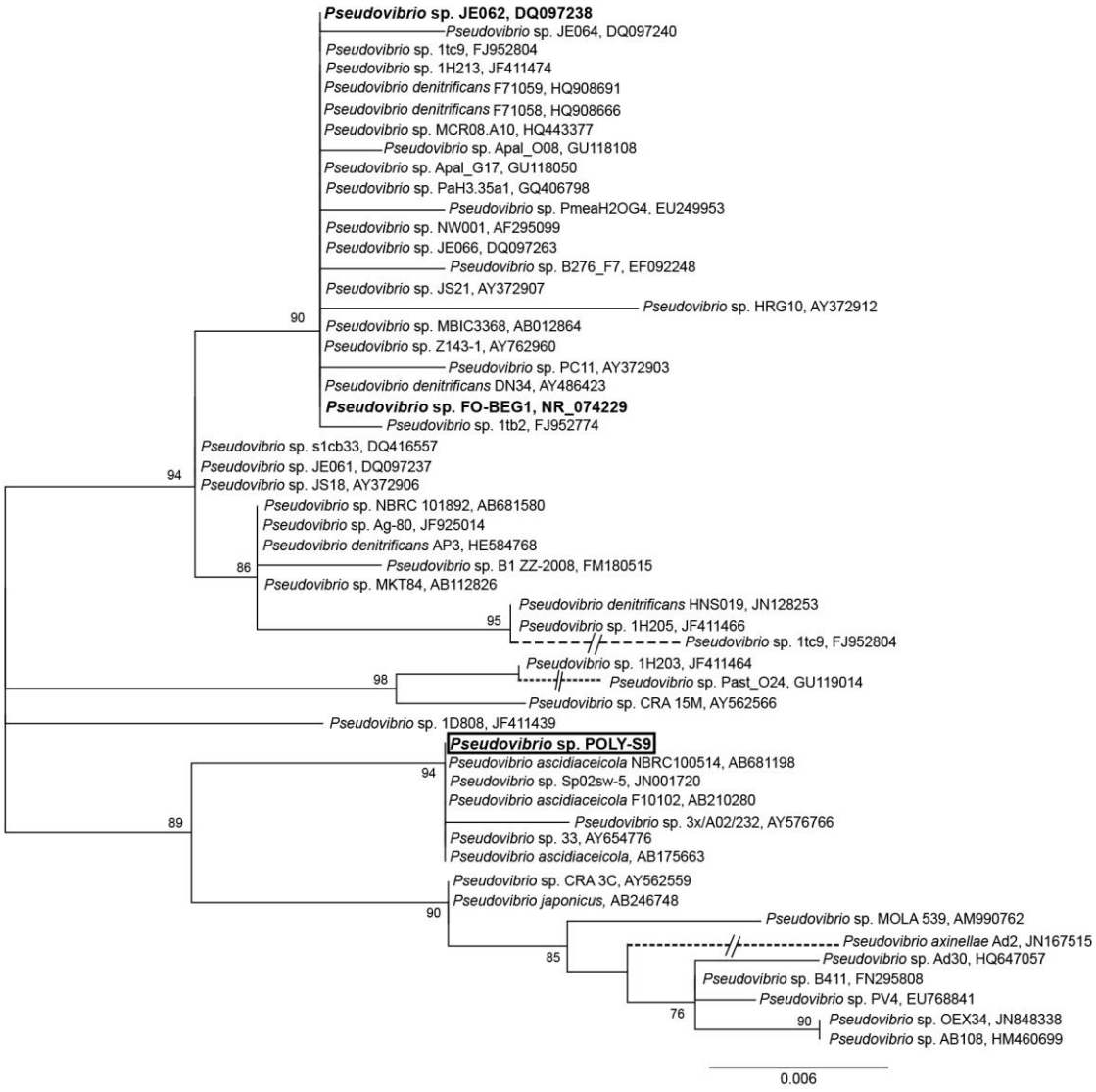


Figure 7- Maximum likelihood tree of *Pseudovibrio sp.* Tree was adapted from Bondarev et al, 2013 and Strains Poly-S9, JCM19062 and JCM12308 were added. *Pseudovibrio* FO1-BEG is referred to as NR 074229, JE062 as DQ097238, Poly-S9 as PPL9.

Carbon and nitrogen metabolic features of the sponge symbiont

The genomic analyses of the sponge symbiont *Pseudovibrio* sp. POLY-S9 revealed the presence of several functional genomic repertoires of carbon and nitrogen metabolic pathways. Detailed study of POLY-S9 genome indicated the possible mechanism of carbohydrate metabolism through the Entner-Doudoroff (ED) pathway, a widespread alternative (Kersters and Ley, 1968) and predominant pathway in marine Alphaproteobacteria (Kumar et al., 2012; Tang et al., 2009). A recent study reported the obligatory role of the ED pathway for gluconate utilization and the pathogenicity in *Vibrio cholerae* (Patra et al., 2012). Consistent with the previously reported *Pseudovibrio*

genomes (FO-BEG1 and JE062), the genes encoding for proteins involved in tricarboxylic acid (TCA) cycle and pentose phosphate pathway were also identified in the current genome. The genes for glycolysis and pentose phosphate pathway have been reported in other sponge -associated microorganisms, *Poribacteria* (Siegl et al., 2011) and *Cenarchaeum symbiosum* (Hallam et al., 2006), indicating a wide range of sugar and carbon utilization mechanism by sponge-symbionts to meet their energy requirements.

In addition, *Pseudovibrio* sp. POLY-S9 genome harbored genes responsible for nitrogen assimilation and denitrification pathways (figure 8). The metabolic waste product of sponges, ammonia, acts as a rich environment for the growth of symbiotic bacteria in nitrogen-deficient oceans (Hentschel et al., 2012). Consistent with previously sequenced *Pseudovibrio* strains FO-BEG1 and JE062, we observed gene clusters responsible for denitrification in POLY-S9 genome. Further analyses of metabolic utilization of the current strain under laboratory conditions could help to understand the detailed physiological traits.

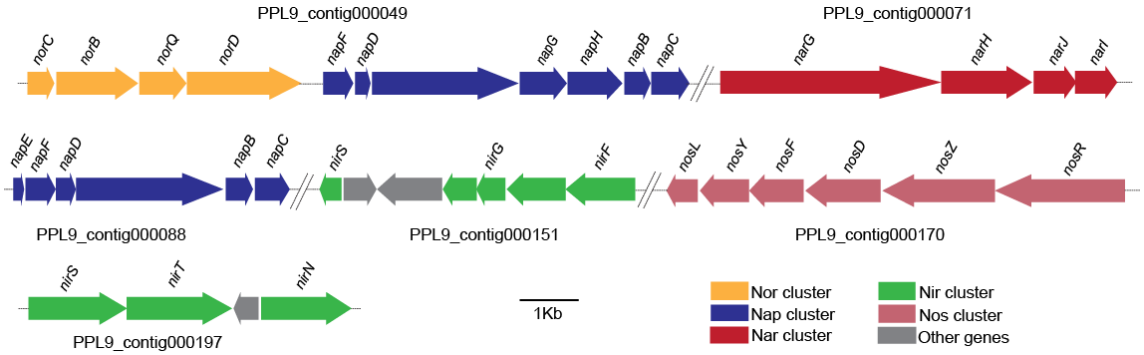
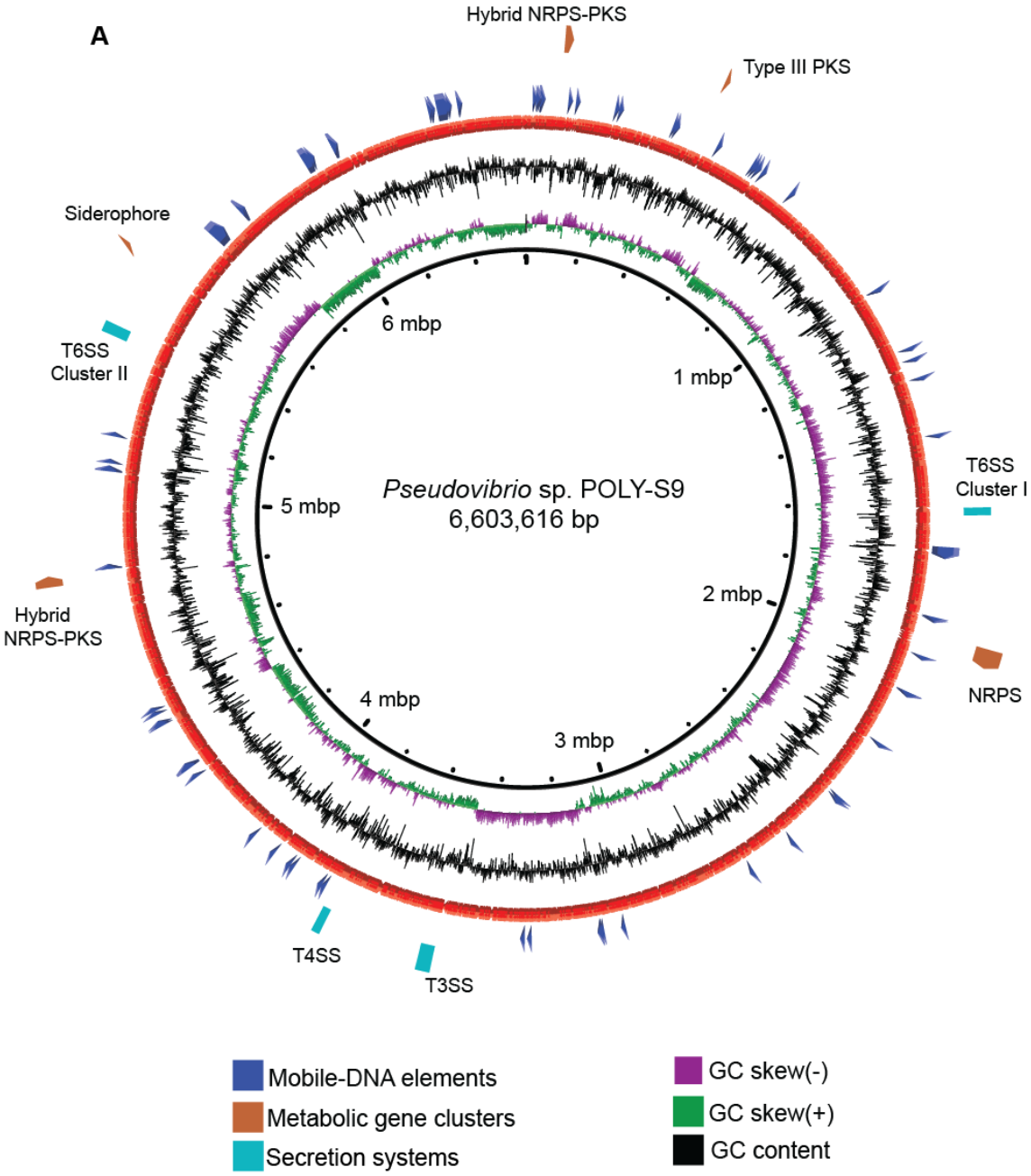


Figure 8 – Physical map of denitrifying gene clusters detected in *Pseudovibrio* sp. POLY-S9 genome. Denitrifying gene clusters (*nor*, *nap*, *nar*, *nir* and *nos*) distributed across the genome in different contigs are shown. Each gene cluster is color coded differently.

Genome representation of General features:



Figures 9 Circular view of *Pseudovibrio* sp. POLY-S9 representing mobile elements and type IV secretion system (T4SS). (A) The blue ring represents the *Pseudovibrio* sp. POLY-S9 chromosome. The black ring represents GC content followed by GC skew (purple/green color). BRIG tool was used to generate the circular view of bacterial chromosome (Alikhan et al. 2011). The distribution of mobile elements and gene transfer agents (GTA) are shown as red clock wise arrows. Brown (outermost ring) block shows the location of T4SS in the genome of POLY-S9.

Table 2- General genome features of *Pseudovibrio sp.* strains used in this work

Pseudovibrio strain:	**Genome size**	**No. Of contigs**	**GC content**	**No of CDS**	**No of rRNA**	**No. Of tRNA**	**Mobile DNA**
FO1-BEG	5916782	2	52.50%	5478	6	69	~2.5%
JE062	57626521	19	52.40%	5225	7	72	~0.3
Poly-S9	6603616	271	51.26%	6171	3	59	~5

Table 3- Cluster of Orthologs Groups (COG) results for each category

COG count			**COG class ID**
FO1-BEG	**JE062**	**Poly-S9**	**--**
3	3	3	B
253	277	280	C
25	24	29	D
440	468	461	E
102	102	99	F
306	345	319	G
182	187	187	H
180	192	211	I
191	197	188	J
387	412	394	K
124	125	165	L
208	216	211	M
85	94	87	N
147	147	158	O
229	247	243	P
118	124	161	Q
540	560	600	R
380	385	429	S
184	207	204	T
78	78	86	U
63	72	62	V

Table 4- Description of the COG categories

COG class ID	Description
B	Chromatin structure and dynamics
C	Energy production and conversion
D	Cell cycle control, cell division, chromosome partitioning
E	Amino acid transport and metabolism
F	Nucleotide transport and metabolism
G	Carbohydrate transport and metabolism
H	Coenzyme transport and metabolism
I	Lipid transport and metabolism
J	Translation, ribosomal structure and biogenesis
K	Transcription
L	Replication, recombination and repair
M	Cell wall/membrane/envelope biogenesis
N	Cell motility
O	Posttranslational modification, protein turnover, chaperones
P	Inorganic ion transport and metabolism
Q	Secondary metabolites biosynthesis, transport and catabolism
R	General function prediction only
S	Function unknown
T	Signal transduction mechanisms
U	Intracellular trafficking, secretion, and vesicular transport
V	Defense mechanisms

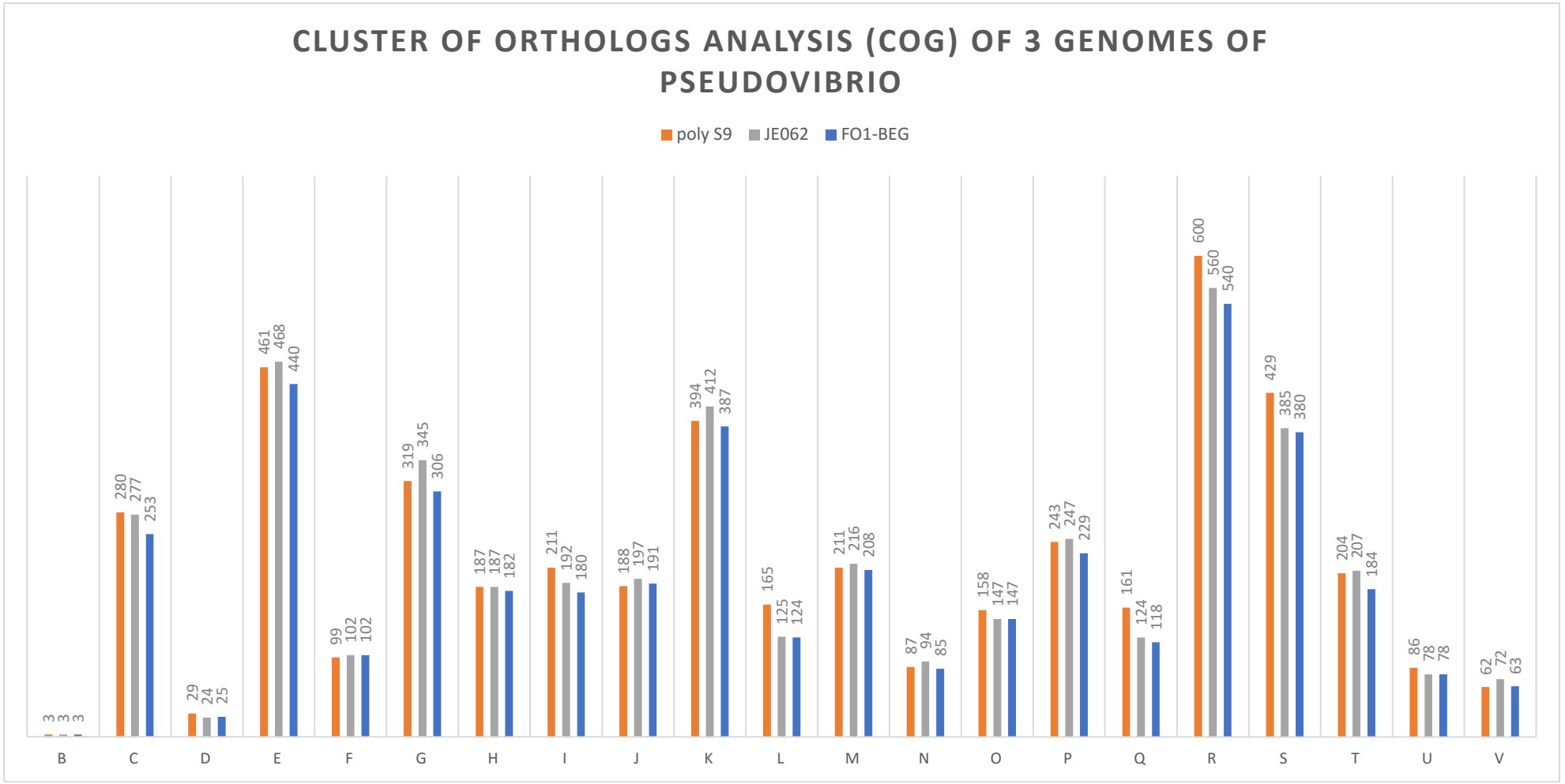


Figure 10 - Graphical representation of the results of the Cluster s of Orthologs Groups (COG). On the x axel we see COG category, as explained on table _ and on the Y axel we see the number of counts for each class. Strain's color identification is shown at the bottom of the graphic

Genomic mining of Secondary Metabolites

Table 5 – Results from the antiSMASH analysis from the 3 genomes

Pseudovibrio strain	Clusters	NRPS	PKS	Hybrid	Other clusters
FO1-BEG	5	×	×	✓ (T3PKS-T1PKS) ✓ (NRPS-T1PKS)	Bacteriocin (x2); Terpene
JE062	3	×	×	✓ (T3PKS-T1PKS)	Bacteriocin; Terpene
Poly-S9	8	✓ ✓	✓ (T3PKS)	✓ (NRPS-T1PKS)	Bacteriocin (x2); Siderophore; Hserlactone

Sponges are proven to be a rich source of bioactive compounds with pharmacological application and in recent years it has been shown that the symbiotic bacteria associated with sponges produce potential novel metabolites (Fuerst, 2014). Most bioactive compounds are produced by polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) (Schwarzer et al., 2003; Staunton and Weissman, 2001). Genomic mining using antiSMASH (Blin et al., 2013) detected several metabolite gene clusters similar to NRPS and PKS in the genome of *Pseudovibrio* sp. POLY-S9, most likely involved in the bioactive compound production. A hybrid NRPS-PKS system (42 kb) similar to *Pseudovibrio* sp. FO-BEG1 strain was observed as a major metabolic gene cluster in POLY-S9 distributed in two different scaffolds (fig.11). According to antiSMASH analyses, this gene cluster is known to be a colibactin biosynthetic gene cluster; common among *E*. *coli* which has previously been described to induce DNA double-strand breaks in eukaryotic cells (Nougayrède et al., 2006). Further investigation of colibactin genomic island by expression and functional analyses could give detailed insight into the biotechnological potential of sponge-associated bacterial hybrid NRPS-PKS system. In addition, POLY-S9 genome coded 5.2 kb type III PKS (locus tag PPL9_00404) exhibiting close similarity with previously reported FO-BEG1 and JE062 polyketide synthase.

Interestingly, a gene cluster coding for NRPS-independent siderophore biosynthetic pathway was detected in POLY-S9 genome. Siderophores are iron chelators synthesized and excreted by many pathogenic/saprophytic microorganisms in response to iron limitation (Drechsel and Jung, 1998). Detailed analyses revealed that these siderophore gene cluster harbored 4 genes (*iucABCD*) coding for aerobactin biosynthesis enzymes (fig. 11). In many virulent *E*. *coli* strains, aerobactin orchestrates iron acquisition, playing a key role in pathogenesis (Dho and Lafont, 1984; Knöbl et al.,

2001; Lafont et al., 1987; Ling et al., 2013). Reanalysis of FO-BEG1 and JE02 genomes using antiSMASH could not detect the presence of siderophore gene clusters. However, PKS/NRPS proteins similar to predicted siderophore biosynthetic gene has been previously reported from various *Pseudovibrio* spp. isolated from Irish marine sponges (O'Halloran et al., 2011) using a PCR screening approach. Whole genome sequencing of *Actinokineospora* sp. strain EG49 isolated from the Red Sea sponge *Spheciospongia vagabunda* also detected the presence of several metabolic gene clusters coding for PKS, NRPS, hybrid NRPS-PKS and siderophore (Harjes et al., 2014) suggesting the need for further investigation for novel bioactive compound from the sponge-associated microbes.

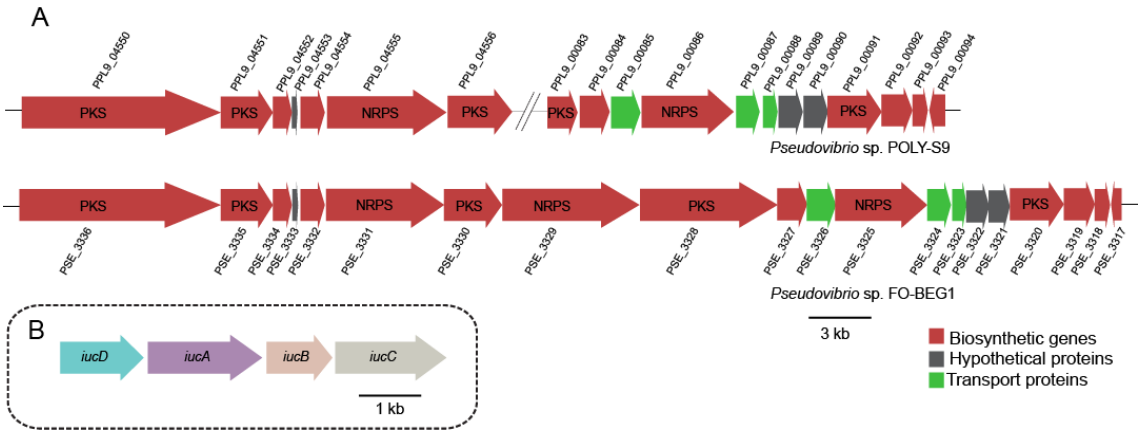


Figure 11- Comparison between metabolic gene clusters of *Pseudovibrio sp.* Poly-S9 and *Pseudovibrio sp.* FO1-BEG. (A) The hybrid non-ribosomal peptide synthetase-peptide synthase (NRPS-PKS) system responsible for colibactin production in the POLY-S9 genome sequenced in this study and the previously sequenced FO-BEG1 genome. Different color codes represent detected ORFs responsible for biosynthesis, hypothetical proteins and transport proteins. Respective locus tags are also labeled. (B) Detected siderophore gene cluster in POLY-S9 coding for aerobactin.

Terpenes, a class of metabolites composed of simple organic compounds were detected in all strains analysed except in the genome of Poly-S9. The term terpene derived from resin component Turpentine. They are usually produce by plants, in particular conifers (Merriam-Webster online dictionary) and other organism such as insects and larvae. Detected terpenes clusters from three genomes are shown in figure 12 (A and B).

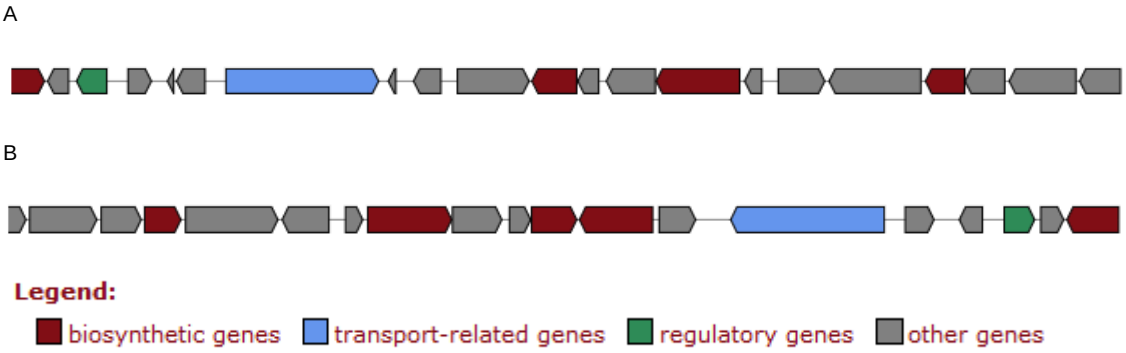


Figure 12- Terpene cluster representation as obtained from the antiSMASH results (A-Fo1-BEG B-JE062)

A particular class of Terpenes, Isocyano terpenes, was first isolated from the sponges in 1973 (Cafieri et al, 1973). To date over 150 structures were characterized. They possess at least one nitrogen atom that forms a cyanide functional group. The mechanism of action of this class of terpenes is yet unclear although it has been tested against common laboratory microbes. The cyanide group present in this class is a known cytotoxin. (Solomonson, 1981). In other cases, some diterpenes have shown to inhibit the growth of some bacteria such as *Bacillus subtlis, Staphylococcus aureus* and *Candida albicans* (Chang et al, 1984; Patra et al, 1984). Also, some antifungal activity has been reported (Fusetani et al, 1990) as well as its toxic effect on *Plasmodium falciparum*, the malaria parasite (Angerhofer and Pezzuto, 1992; König et al., 1996; Wright et al., 1996, 2001; Simpson et al., 1997).

Genes in the COG 'Q' category are overrepresented in *Pseudovibrio* POLY-S9 when compared to FO-BEG1 and JE062 (161, 118, and 124 genes, respectively) (fig.10). The *Pseudovibrio* sp. POLY-S9 isolated from the marine sponge *P. penicillus* may produce novel secondary metabolites and future studies could reveal its potential as a source of bioactive compounds.

Secretion systems

Successful colonization of the host depends on the ability of microbes to interact with the host. Effector molecules such as toxins, proteins and virulence factors mediate these processes and facilitate to control of host cells. Gram negative bacteria use specialized systems, secretion systems (SSs) in order to facilitate the invasion (Cossart and Sansonetti, 2004). Figure 13 shows a representation of the secretion systems, from type 1 (T1SS) to type 6 (T6SS)

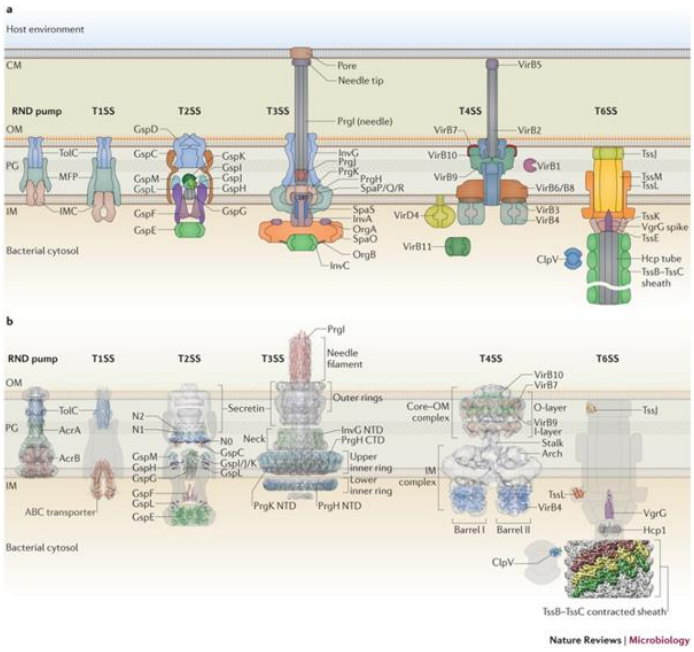


Figure 13- Representation of various types of secretion systems. Adapted from Nature Reviews Microbiology 13, 343–359 (2015))

Type III (T3SS) and type VI (T6SS) (fig. 14 and fig.15) secretion systems were evident in the genome of *Pseudovibrio* sp. POLY-S9 indicating the possible mechanism mediating the interaction between the bacteria and the eukaryotic sponge-hosts.

T3SS or 'injectisome' have been reported to serve the Gram-negative bacteria a central role in pathogenicity/symbiosis (Dale and Moran, 2006; Preston, 2007) by directing the assembly of flagella, secretion of extracellular protein and injecting the effector molecules into the target cells. The exported effector molecules play a major role like invasion and/or evading the host defense system that enhances the bacterial survival within the host. Apart from T3SS structural component genes, we found the homologues of three effector molecules – two copies of YpkA (*Yersenia* protein kinase A), one copy each of YopJ (*Yersenia* outer protein J) and IpgD in POLY-S9 genome (fig. 14). YpkA is known to disrupt the actin-based cytoskeletal system, thus interfering with the phagocytotic activity (Wiley et al., 2006, 2009). Whereas, IpgD and YopJ causes cytoskeleton rearrangement at the host-entry site after invading the host cell (Niebuhr et al., 2000) and inhibition of the host innate immune response (Mukherjee et al., 2006). Current genome analyses confirm the role of effector proteins in mediating the colonization of *Pseudovibrio* sp. POLY-S9 with the eukaryotic sponge-host and successful establishment of symbiosis by evading the phagocytosis.

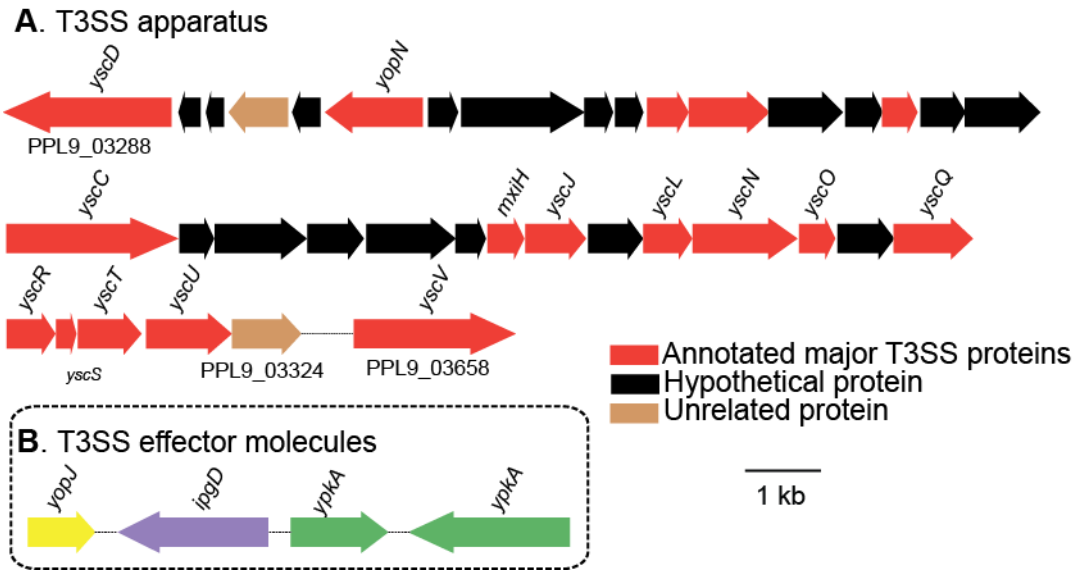


Figure 14- . Genetic organization of type III secretion system (T3SS) and its effector molecules in *Pseudovibrio* sp. POLY-S9. (A) Major genes annotated for T3SS structural components are colored as red blocks. Black and brown blocks represents hypothetical and unrelated genes, respectively. (B) Effector molecules *yopJ*, *ipgD* and *ypkA* hololougues responsible for the distruption of host cell mechanism.

We detected T6SS, a complex multi-component secretion machine often involved in interaction between the hosts and the pathogenic/symbiotic bacteria (Filloux et al., 2008). This widespread secretion system represents an alternative protein translocation pathway (Mougous et al., 2006), enabling the pathogenic bacteria to outcompete other commensal bacteria (Kapitein and Mogk, 2013), biofilm formation (Aschtgen et al., 2008), quorum sensing (Sheng et al., 2013) and antipathogenesis (Jani and Cotter, 2010). Genome of POLY-S9 encoded two gene clusters representing major T6SS core proteins (fig. 15). Multiple copies of genes coding for type VI secretion system effector proteins, *hcpl* (3 copies) and *vgrG* (7 copies) were also evident in other regions of the chromosome (fig.15). Components of this newly described secretion system have been partially characterized. However, T6SS is known to be a key virulence factor for a few pathogenic bacteria (*Vibrio cholera*, *Rhizobium leguminosarum*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *E. coli*, *Burkholderia mallei*) and implicated in transport of proteins into the eukaryotic hosts (Bingle et al., 2008). It has been reported that T6SSs not only establish the pathogenicity but also function to promote the commensal/mutualistic association with the host (Jani and Cotter, 2010).

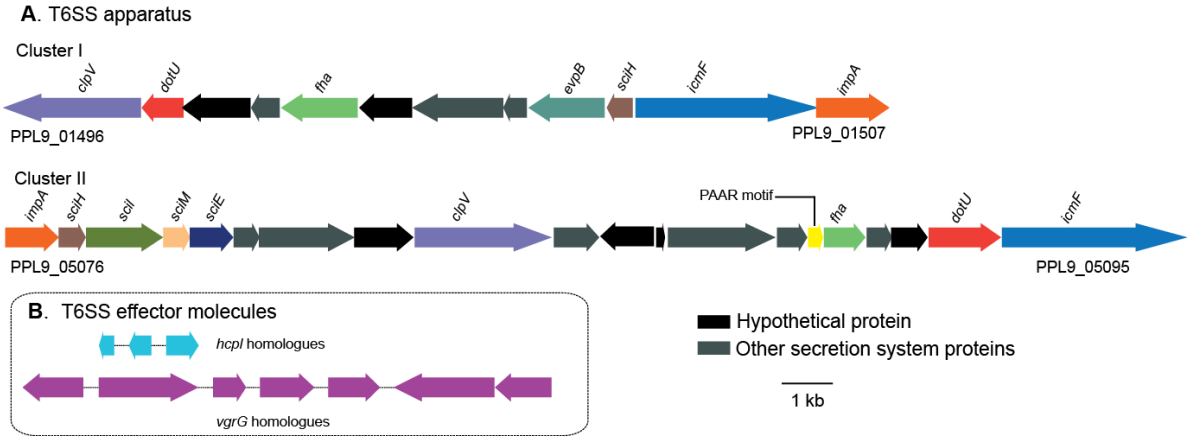


Figure 15 - T6SS apparatus Genetic organization of type VI secretion system (T6SS) gene clusters and its effector molecules in *Pseudovibrio* sp. POLY-S9**.** (A) Genes coding for T6SS apparatus are distributed in two clusters. Blocks of related genes are represented in the same color. Hypothetical and other secretion system related protein-coding genes are colored black and gray, respectively. (B) Genes coding for T6SS effector molecules (3 copies of *hcpI* and 7 copies of *vgrG*).

Type IV secretion system and pathogenicity

The presence of type IV secretion system (T4SS) protein complexes was a distinct feature of the sponge-associated bacterium, *Pseudovibrio* sp. POLY-S9. T4SSs are involved in mediating the DNA transfer through conjugation, DNA uptake/release from the extracellular milieu and translocation of virulence/effector proteins into the host cells (Alvarez-Martinez and Christie, 2009). T4SSs loci of *Pseudovibrio* sp. POLY-S9 encoded the proteins homologous to several virulence genes, *virB2*, *virB4*, *virB6*, *virB9*, *virB10*, *virB11* and *virD4* (fig. 16). These protein subunits of T4SSs forms the pilus and the translocation channels spanning the bacterial cell envelop (Voth et al., 2012). Besides *vir* genes, the T4SSs loci also encoded the genes responsible for the conjugation (origin of transfer (*oriT*) and relaxase). Relaxase-*oriT* complex is required to initiate the lateral transfer and incorporation of large segment of DNA (genomic island) islands into the bacterial chromosome (Waldor, 2010).

In clinical microbes like *Haemophilus* and *Pseudomonas*, T4SSs are also utilized to mediate the horizontal gene transfer (HGT) augmenting a significant role in the evolution of several pathogenic characters like antibiotic-resistance and virulence (Juhas, 2015). The versatile nature of detected T4SSs in *Pseudovibrio* sp. POLY-S9 might result in the genome plasticity, enabling the bacterium to adapt to various niches in response to changes in their environment.

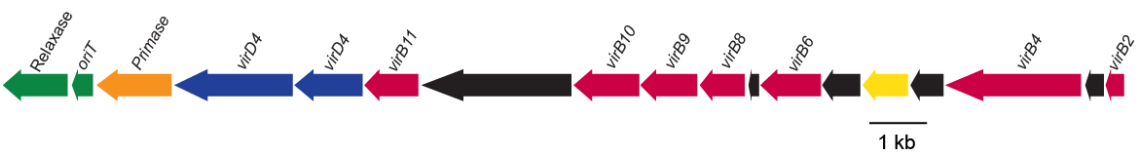


Figure 16 - Genetic organization of T4SS detected in the genome of Pseudovibrio sp. POLY-S9.

Table 6- Gene, gene product and functional category of the detected components of the secretion systems found in Poly-S9 genome

	Gene Product	Gene	Functional category
Type III secretion systems (T3SSs)	Yop protein translocation protein D	*yscD*	Involved in flagellar assembly, secretion of extracellular proteins and effector molecules into the host cell
	Hypothetical protein		
	Hypothetical protein		
	Cyclic nucleotide-binding domain protein		
	Hypothetical protein		
	Membrane-bound Yop targeting protein YopN	*yopN*	
	Hypothetical protein		
	Hypothetical protein		
	Hypothetical protein		
	Hypothetical protein		
	Tetratricopeptide repeat protein		
	Secretion system effector C (SsecC) like family		
	Hypothetical protein		
	Hypothetical protein		
	Tetratricopeptide repeat protein		
	Hypothetical protein		
	Hypothetical protein		
	Yop protein translocation protein C	*yscC*	
	Hypothetical protein		
	Hypothetical protein		
	Hypothetical protein		
	Hypothetical protein		
	Hypothetical protein		
	Type III secretion needle MxiH like protein	*mxiH*	
	Yop proteins translocation J	*yscJ*	
	Hypothetical protein		
	Yop proteins translocation protein L	*yscL*	
	ATP synthase yscN (Yop proteins secretion ATPase)	*yscN*	
	Type III secretion protein YscO		
	Hypothetical protein		
	Yop proteins translocation protein Q	*yscQ*	
	Yop proteins translocation protein R	*yscR*	
	Yop proteins translocation protein S	*yscS*	
	Yop proteins translocation protein T	*yscT*	
	Yop proteins translocation protein U	*yscU*	
	Dienelactone hydrolase-like protein		
	Low calcium response locus protein D	*yscV*	
T3SS	Effector protein YopJ (Virulence factor YopJ)	*yopJ*	Responsible for the disruption of host

	Inositol phosphate phosphatase ipgD (effector protein ipgD)	*ipgD*	cytoskeletal system and phagocytotic activity
	Protein Kinase YpkA	*ypkA*	
	Protein Kinase YpkA	*ypkA*	
	Gene Product	**Gene**	**Functional category**
Type VI secretion systems (T6SSs) Cluster I	ATPase, type VI secretion system ClpV1	*clpV1*	Widespread secretion system responsible for protein translocation, enhancing the pathogenicity and biofilm formation
	Type IV / VI secretion system DotU	*dotU*	
	Hypothetical protein		
	Type VI secretion lipoprotein		
	FHA domain protein	*fha*	
	Putatative protein		
	Type VI secretion system		
	Type VI secretion system, lysozyme-related protein		
	type VI secretion protein EvpB	*evpB*	
	Cytoplasmic protein SciH	*sciH*	
	Type VI secretion protein IcmF	*icmF*	
	ImpA domain protein	*impA*	
Type VI secretion systems (T6SSs) Cluster II	ImpA domain protein	*impA*	
	Cytoplasmic protein SciH	*sciH*	
	Cytoplasmic protein SciI	*SciI*	
	SciM protein	*SciM*	
	Virulence protein, SciE type	*sciE*	
	Type VI secretion system, lyzozyme-related protein		
	Type VI secretion system		
	Putative protein		
	Type VI secretion ATPase, ClpVI family	*clpV*	
	Voltage-gated sodium channel		
	Hypothetical protein		
	Hypothetical protein		
	Rhs element Vgr protein		
	Hypothetical protein		
	PAAR motif protein		
	FHA domain protein	*fha*	
	Type VI secretion lipoprotein		
	Hypothetical protein		
	Type IV/VI secretion system, DotU/OmpA/MotB	*dotU*	
	Type VI secretion protein IcmF	*icmF*	
T6SS effector proteins	Type VI secretion system effector, Hcp1	*hcpI*	Major virulence factors for pathogenic bacteria
	Type VI secretion system effector, Hcp1	*hcpI*	
	Type VI secretion system effector, Hcp1	*hcpI*	
	Type VI secretion system Vgr family protein	*vgrG*	
	Type VI secretion system Vgr family protein	*vgrG*	
	Type VI secretion system Vgr family protein	*vgrG*	

	Type VI secretion system Vgr family protein	*vgrG*	
	Type VI secretion system Vgr family protein	*vgrG*	
	Type VI secretion system Vgr family protein	*vgrG*	
	Type VI secretion system Vgr family protein	*vgrG*	

	Gene Product	**Gene**	**Functional category**
Type IV secretion systems (T4SSs)	Relaxase/mobilization nuclease domain protein		Secretion system protein involved in conjugation, uptake of foreign DNA from the surrounding and transport of virulence factors into the host cell
	OriT binding protein	*oriT*	
	DNA primase		
	VirD4 protein	*virD4*	
	VirD4 protein	*virD4*	
	VirB11 protein	*virB11*	
	Hypothetical protein		
	VirB10 protein	*virB10*	
	VirB9 protein	*virB9*	
	VirB8 protein	*virB8*	
	Hypothetical protein		
	VirB6 protein	*virB6*	
	Hypothetical protein		
	GTP-binding protein		
	Hypothetical protein		
	VirB4 protein	*virB4*	
	Hypothetical protein		
	VirB2 family protein	*virB2*	
Quorum sensing (QS) system - luxR/luxI transcriptional regulators	transcritptional regulator, LuxR family protein	*luxR*	Bacterial communication and regulation of bacterial processes such as gene expression in response to cell population density, biofilm formation, symbioses etc
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	transcritptional regulator, LuxR family protein	*luxR*	
	Autoinducer synthatase	*luxI*	
***tad* (tight adherence) locus**	type II/IV secretion system protein, TadA subfamily	*cpaF*	Genes essential for adherence, biofilm formation, synthesis of Flp pili and pathogenesis
	response regulator receiver protein	*cpaE*	
	pilus assembly protein CpaD	*cpaD*	
	bacterial type II/III secretion system protein	*cpaC*	
	pilus assembly, Flp-type CpaB	*cpaB*	

	Gene Product	Gene	Functional category
Invasion associated locus	Invasion associated locus B	*ialB*	Virulence factors mediating host-microbe association
	Invasion associated locus B	*ialB*	
	Invasion associated locus B	*ialB*	
	Invasion associated locus B	*ialB*	
	Invasion associated locus B	*ialB*	
	Invasion associated locus B	*ialB*	
	Invasion associated locus B (IalB) protein	*ialB*	
	Invasion associated locus B	*ialB*	
	Invasion associated locus B	*ialB*	
Amyloid production genes	Curli production assembly	*csgF*	Gene cluster responsible for curli fiber production and assembly. Mediates microbial attachment to the eukaryotic cell surface
	Curli production assembly	*csgG*	
	Hypothetical		
	Curlin associated repeat protein		
	Curlin associated repeat protein		
	Curlin associated repeat protein		
YadA/TadE like protein	protein containing YadA-like, C-terminal domain		Adhesion factors
	protein containing YadA-like, C-terminal domain		
	TadE-like protein	*tadE*	Colonization factors
	TadE-like protein	*tadE*	
NRPS-PKS hybrid gene cluster	Polyketide synthase		A major metabolic gene cluster known for the biosynthesis of colibactin -resposible to induce DNA double-strand break in eukaryotic cells
	Polyketide synthase		
	3-hydroxybutyryl-CoA dehydrogenase	*hbd*	
	Hypothetical protein		
	Acyl-CoA dehydrogenase domain protein		
	Non-ribosoma peptide synthetase		
	Polyketide synthase		
	Putative peptide synthetase		
	Asp-tRNAAsn/Glu-tRNAGln amidotransferase A		
	Multi antimicrobial extrusion protein MatE		
	Non-ribosoma peptide synthetase		
	Secretion protein HlyD	*hlyD*	
	Lipoprotein-releasing system ATP-binding protein LolD	*lolD*	
	protein containing DUF214, permase predicted		
	protein containing DUF214, permase predicted		
	Polyketide synthase		
	Beta-lactamase class C		
	Cadicidin biosynthesis thioesterase		
	4'-phosphopantetheinyl transferase	*hetI*	

Eukaryotic-like proteins that facilitate invasion and colonization from microorganisms

In the genome of *Pseudovibrio* sp. POLY-S9, several genes were identified encoding for eukaryotic-like proteins (ELPs), which contain motifs such as ankyrin-repeats (ANKs), tetratrico peptide repeats (TPRs) and Sel1 repeats (table 6). These symbiosis factors (eukaryotic-like proteins) are widely represented in the genomes of pathogenic as well as symbiotic microbes, and postulated to mediate the host behavior by interfering eukaryotic protein-protein interactions. Sponge-associated microbes have been frequently reported to encode ELPs. For instance, metagenomic and metaproteogenomic analysis of *C*. *concentrica* revealed the presence of ANKs and TRPs (Liu et al., 2012; Thomas et al., 2010), and in several other marine sponge-associated microbial communities (Fan et al., 2012). Whole genome analyses of sponge symbionts, *Poribacteria* (Siegl et al., 2011), *Deltaproteobacteria* (Liu et al., 2011) and *Pseudovibrio* strains (Bondarev et al., 2013) also showed the presence of eukaryotic-domain containing motifs. Other than sponge-symbionts, the abundance of ANKs encoding genes have been reported in the genomes of obligate and facultative symbionts such as *Wolbachia pipientis* (Iturbe-Ormaetxe et al., 2005), *Ehrlichia canis* (Mavromatis et al., 2006), *Legionella pneumophila* (Habyarimana et al., 2008) and *Coxiella burnetii* (Voth et al., 2009). It is well understood that the secretion of ANK-containing protein by pathogenic bacteria *L*. *pneumophila* interfere with the eukaryotic host cell functions such as polyubiquitylation (Al-Khodor et al., 2008) and vesicular transport (Pan et al., 2008). Whereas, ANK of *C*. *burnetti* facilitate the survival of the bacteria in the mammalian host by preventing pathogen induced apoptosis (Lührmann et al., 2010). The possible mechanism of symbiotic association between *Pseudovibrio* sp. POLY-S9 and the sponge hosts was further confirmed by the presence of TPR and Sel1 repeats. These repeat-containing proteins might play a role in establishing the interaction by – allowing the bacterial entry into the host cell, regulating the exopolysaccharide synthesis (Mittl and Schneider-Brachert, 2007) or regulating the intracellular trafficking (Schmitz-Esser et al., 2010). The enrichment of *Pseudovibrio* sp. POLY-S9 with various ELPs further affirms the possible mechanism mediating the symbiotic association with the intertidal sponge *P*. *penicillus*.

Adhesion and invasion factors

Consistent with the previously sequenced *Pseudovibrio* strains, we detected a gene cluster responsible for amyloid (curli) production and assembly/transport components, and eight homologues of the *Bartonella bacilliformis ialB* (invasion-associated locus B) in the genome of POLY-S9 (table 6). The function of curli fibres are assumed to be related to bacterial adhesion mechanism mediating attachment and biofilm formation (Kikuchi et al., 2005). Curli systems are widespread in bacteria from diverse habitats (Dueholm et al., 2012) and are considered to act as a virulence factor by mediating the interaction with a wide range of host protein such as extracellular matrix protein (Collinson et al., 1993; Olsén et al., 1989). These adhesion factors could play an important role in symbiosis too. For instance, many members of Rhizobiales which form symbiotic association with the root nodules and plant-associated *Salmonella enterica* (Barak et al., 2005), has been reported to encode curli genes. It is believed that *Pseudovibrio* sp. might mediate the adhesion through curli fibre production and assembly. Invasion-associated locus B (*ialB*) gene product are known to be a major virulence factor in *B. bacilliformis* directing the erythrocyte parasitism (Rolain et al., 2002). We also detected the genes coded for proteins containing YadA (*Yersenia* adhesin A) and TadE-like domains in the genome of POLY-S9. These are major cell adhesion factors crucial in determining the virulence of the pathogenic bacteria crucial for extracellular matrix-specific cell adhesion (Heise and Dersch, 2006) and colonization (Schreiner et al., 2003).

A widespread genomic island consisting of *tad* (tight adherence) locus (*cpaABCDEF*) was detected in the POLY-S9 genome (fig. 17). The *tad* genes are essential for adherence, biofilm formation, synthesis of bundled Flp (fimbrial low-molecular-weight) pili and pathogenesis in a wide variety of Gram-positive and Gram-negative bacteria (Tomich et al., 2007). Phylogenetic studies indicate that *tad* locus has undergone a series of gene duplication, gene loss, recombination and horizontal gene transfer (HGT) between distant bacterial relatives (Filloux, 2010; Planet et al., 2003). Recurrent detection of *tad* locus in the current genome and previously sequenced *Pseudovibrio* strains emphasizes the evolutionary significance of widespread colonization island in establishing the bacterial association with the eukaryotic sponge hosts inhabiting different environmental niches and its possible procurement through HGT as an adaptation to colonize a wide range of hosts.

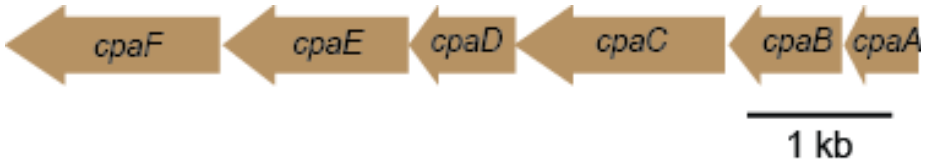


Figure 17- Physical map of *tad* locus detected in the genome of *Pseudovibrio* sp. POLY-S9. The genes are represented inside the blocks.

Interaction among sponge-associated bacteria

Bacterial communication, an essential aspect in a community where microbes compete for resource is achieved by a mechanism termed quorum sensing (QS). QS system seems to regulate bacterial processes such as biofilm formation, virulence factor secretion, bioluminescence, motility, antibiotic production, sporulation and DNA uptake (Ng and Bassler, 2009). It is induced by diffusible signaling molecules, autoinducers (transcriptional regulators) which triggers expression of specific genes (Fuqua et al., 1994). The genome analysis of strain POLY-S9 revealed the presence of proteins containing DNA binding LuxR domains, where they act as transcriptional regulators of N-acyl homeserine lactones (AHL), a common proteobacterial signaling molecules. Genome analyses of *Pseudovibrio* strains FO-BEG and JE062 reported the presence of LuxR, and absence of genes coding for the AHL synthase, hypothesizing that these strains respond to the autoinducers secreted by other bacterial species (Case et al., 2008) or communicate in an autoinducer-independent way (Patankar and González, 2009). However, we found *luxI* homologue coding for AHL synthase in the genome of *Pseudovibrio* POLY-S9 suggesting its ability to synthesize quorum sensing molecules and communication within their own species.

Mobile elements and Gene transfer agents

When compared to previously sequenced genomes of the *Pseudovibrio* strains isolated from the scleractinian coral and the sponge *M*. *laxissima* (Bondarev et al., 2013), it is evident that the *Pseudovibrio* sp. POLY-S9 strain has a higher number of prophage genes, a probable reason to explain its genome size expansion (fig. 9). It is known that the facultative intracellular bacteria usually harbour four-fold more mobile-DNA than the obligate intracellular bacteria (Bordenstein and Reznikoff, 2005), which is consistent with the predictions that the free-living bacteria tend to have more mobile-DNA content similar to facultative ones. As evident from other intracellular bacteria, transposable elements constitute the major portion of mobile-DNA in the *Pseudovibrio* sp. POLY-S9 strain. In contrast, sponge symbiont *Pseudovibrio* sp. FO-BEG1 and vertically transmitted

symbiont *Pseudovibrio* sp. JE062 strains (Bondarev et al., 2013) harbored less mobile elements. The lifestyle of a bacterium might influence the genome plasticity, in which the bacteria with small population size (obligate bacteria) might experience evolutionary processes resulting in the loss of mobile-DNA (Bordenstein and Reznikoff, 2005). The large amount of mobile elements detected in the genome of *Pseudovibrio* sp. POLY-S9 strain indicates its permissive lifestyle capable of host-switching.

We observed several set of gene transfer agents (GTAs) in the *Pseudovibrio* sp. POLY-S9 genome. GTAs (fig. 9), phage-like entities are responsible for horizontal gene transfer (HGT) by packaging and transferring random segments of host bacterial DNA to a recipient cell of closely or distantly related lineages (Zhaxybayeva and Doolittle, 2011). It is hypothesized that the selection drives the maintenance of GTAs and thus enhances the HGT process, facilitating the adaptive evolution of the host-adaptation systems and allowing the host-range size expansion (Guy et al., 2013).

Conclusions:

The microbial association plays an important role in order to maintain the sponge health and survival by providing beneficial supplements. Due to the limitations in the isolating the symbiotic bacteria, still a vast majority of microbes and its potential and the evolutionary genomics of symbioses are not known to us. For instance, the role of symbiotic bacteria for nitrogen fixation has been suggested using the molecular techniques by searching for the nitrogen fixing gene, *nifH* (Mohamed et al. 2008) from the whole bacterial communities. However, the advancement in sequencing technique made the sponge-symbioses research approachable (Liu et al., 2012; Thomas et al., 2010) and enabled to gain knowledge regarding the mechanisms underlying the sponge-symbiont interactions.

Genomic analysis performed during this work allowed us to have a better insight into the association of bacteria with its eukaryotic sponge and the molecular mechanisms involved in lifestyle adaptation to specific habitat. The genomic flexibility for the symbiotic lifestyle includes their potential ability to interact with the invertebrate host and symbioses. This is further confirmed by the presence of genes responsible for the adhesion molecules, T3SSs, T4SSs and T6SSs playing an important role in initiating the cell surface contact and the secretion of effector molecules into the host cells. In addition, the presence of genes coding for the eukaryotic-like repeat domains (ankyrin-repeat domains, tetratrico peptide repeat domains, Sel1, TadE like domains) function as effector molecules, which interfere with the host's immune system. *Pseudovibrio sp.* Poly-S9 also exhibited a genomic repertoire containing several bioactive gene clusters, namely T1

and T3PKS and Siderophore and the NRPS clusters. Future analyzes of these gene cluster by heterologous gene expression may help to elucidate the bioactive activity. The high percentage of mobile elements in this genome may confer a better adaptation capacity to a variety of habitats and hosts by facilitating the exchange of genetic material through horizontal gene transfer. This study concluded the host-adaptability features of Pseudovibrio sp. POLY-S9 to establish symbiotic association with the marine sponges.

The work done with this bacterial genome may represent a foundation work for future investigation regarding microbes that live in association with intertidal marine sponges. In future, this study may be broadened by incorporating more genomes of the symbiotic microbes. This work provided me with a new insight into the world of Bioinformatics. As a bachelor student in Biochemistry I was fortunate to participate in the study that served as the foundation of this work by being part of all stages, from sampling of the marine sponges and isolation of the sponge-associated microbes. However, the present work proven to be very different from my previous research activities. During the bachelor dissertation, I was engaged with more practical work, consisting of culturing, isolation of the microbes, extraction, purification of the genetic material, and genetic identification of the isolates based on molecular markers like 16S rRNA genes. As referred before, being a Biochemistry student, my training during bachelor course allowed me to conduct and troubleshoot the research problem while pursuing master thesis.

References

Agogué, H., Casamayor, E.O., Bourrain, M., Obernosterer, I., Joux, F., Herndl, G.J., and Lebaron, P. (2005). A survey on bacteria inhabiting the sea surface microlayer of coastal ecosystems. FEMS Microbiol. Ecol. *54*, 269–280.

Alex, A., and Antunes, A. (2015). Pyrosequencing Characterization of the Microbiota from Atlantic Intertidal Marine Sponges Reveals High Microbial Diversity and the Lack of Co-Occurrence Patterns. PLoS ONE *10*.

Al-Khodor, S., Price, C.T., Habyarimana, F., Kalia, A., and Abu Kwaik, Y. (2008). A Dot/Icm-translocated ankyrin protein of Legionella pneumophila is required for intracellular proliferation within human macrophages and protozoa. Mol. Microbiol. *70*, 908–923.

Alvarez-Martinez, C.E., and Christie, P.J. (2009). Biological diversity of prokaryotic type IV secretion systems. Microbiol. Mol. Biol. Rev. MMBR *73*, 775–808.

Aschtgen, M.-S., Bernard, C.S., De Bentzmann, S., Lloubès, R., and Cascales, E. (2008). SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative Escherichia coli. J. Bacteriol. *190*, 7523–7531.

Attaway, D.H., and Zaborsky, O.R. (1993). Pharmaceutical and Bioactive Natural Products (Springer US).

Barak, J.D., Gorski, L., Naraghi-Arani, P., and Charkowski, A.O. (2005). Salmonella enterica virulence genes are required for bacterial attachment to plant tissue. Appl. Environ. Microbiol. *71*, 5685–5691.

Bingle, L.E., Bailey, C.M., and Pallen, M.J. (2008). Type VI secretion: a beginner's guide. Curr. Opin. Microbiol. *11*, 3–8.

Blin, K., Medema, M.H., Kazempour, D., Fischbach, M.A., Breitling, R., Takano, E., and Weber, T. (2013). antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. *41*, W204–W212.

Blunt, J.W., Copp, B.R., Munro, M.H., Northcote, P.T., and Prinsep, M.R. (2006). Marine natural products.

Blunt, J.W., Copp, B.R., Hu, W.-P., Munro, M.H., Northcote, P.T., and Prinsep, M.R. (2007). Marine natural products. Nat. Prod. Rep. *24*, 31–86.

Bondarev, V., Richter, M., Romano, S., Piel, J., Schwedt, A., and Schulz-Vogt, H.N. (2013). The genus Pseudovibrio contains metabolically versatile bacteria adapted for symbiosis. Environ. Microbiol. 2095–2113.

Bordenstein, S.R., and Reznikoff, W.S. (2005). Mobile DNA in obligate intracellular bacteria. Nat. Rev. Microbiol. *3*, 688–699.

Case, R.J., Labbate, M., and Kjelleberg, S. (2008). AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. ISME J. *2*, 345–349.

Chiu, J.M.Y., Li, S., Li, A., Po, B., Zhang, R., Shin, P.K.S., and Qiu, J.-W. (2012). Bacteria associated with skeletal tissue growth anomalies in the coral Platygyra carnosus. FEMS Microbiol. Ecol. *79*, 380–391.

Collinson, S.K., Doig, P.C., Doran, J.L., Clouthier, S., Trust, T.J., and Kay, W.W. (1993). Thin, aggregative fimbriae mediate binding of Salmonella enteritidis to fibronectin. J. Bacteriol. *175*, 12–18.

Dale, C., and Moran, N.A. (2006). Molecular interactions between bacterial symbionts and their hosts. Cell *126*, 453–465.

Dho, M., and Lafont, J.P. (1984). Adhesive properties and iron uptake ability in Escherichia coli lethal and nonlethal for chicks. Avian Dis. *28*, 1016–1025.

Diaz, M.C., and Rützler, K. (2001). Sponges: an essential component of Caribbean coral reefs. Bull. Mar. Sci. *69*, 535–546.

Donia, M.S., Hathaway, B.J., Sudek, S., Haygood, M.G., Rosovitz, M.J., Ravel, J., and Schmidt, E.W. (2006). Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. Nat. Chem. Biol. *2*, 729–735.

Donia, M.S., Ravel, J., and Schmidt, E.W. (2008). A global assembly line to cyanobactins. Nat. Chem. Biol. *4*, 341–343.

Donia, M.S., Ruffner, D.E., Cao, S., and Schmidt, E.W. (2011a). Accessing the Hidden Majority of Marine Natural Products Through Metagenomics. Chembiochem *12*, 1230–1236.

Donia, M.S., Fricke, W.F., Ravel, J., and Schmidt, E.W. (2011b). Variation in Tropical Reef Symbiont Metagenomes Defined by Secondary Metabolism. PLoS ONE *6*.

Donia, M.S., Fricke, W.F., Partensky, F., Cox, J., Elshahawi, S.I., White, J.R., Phillippy, A.M., Schatz, M.C., Piel, J., Haygood, M.G., et al. (2011c). Complex microbiome underlying secondary and primary metabolism in the tunicate-Prochloron symbiosis. Proc. Natl. Acad. Sci. U. S. A. *108*, E1423–E1432.

Drechsel, H., and Jung, G. (1998). Peptide siderophores. J. Pept. Sci. *4*, 147–181.

Dueholm, M.S., Albertsen, M., Otzen, D., and Nielsen, P.H. (2012). Curli Functional Amyloid Systems Are Phylogenetically Widespread and Display Large Diversity in Operon and Protein Structure. PLoS ONE *7*, e51274.

Enticknap, J.J., Kelly, M., Peraud, O., and Hill, R.T. (2006). Characterization of a Culturable Alphaproteobacterial Symbiont Common to Many Marine Sponges and

Evidence for Vertical Transmission via Sponge Larvae. Appl. Environ. Microbiol. *72*, 3724–3732.

Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N.S., and Thomas, T. (2012). Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. Proc. Natl. Acad. Sci. *109*, E1878–E1887.

Faulkner, D.J. (2000). Marine natural products. Nat. Prod. Rep. *17*, 7–55.

Faulkner, D.J. (2001). Marine natural products. Nat. Prod. Rep. *18*, 1R – 49R.

Filloux, A. (2010). A Variety of Bacterial Pili Involved in Horizontal Gene Transfer. J. Bacteriol. *192*, 3243–3245.

Filloux, A., Hachani, A., and Bleves, S. (2008). The bacterial type VI secretion machine: yet another player for protein transport across membranes. Microbiol. Read. Engl. *154*, 1570–1583.

Flemer, B., Kennedy, J., Margassery, L.M., Morrissey, J.P., O'Gara, F., and Dobson, A.D.W. (2012). Diversity and antimicrobial activities of microbes from two Irish marine sponges, Suberites carnosus and Leucosolenia sp. J. Appl. Microbiol. *112*, 289–301.

Fondo, E.N., and Martens, E.E. (1998). Effects of mangrove deforestation on macrofaunal densities, Gazi Bay, Kenya. Mangroves Salt Marshes *2*, 75–83.

Fuerst, J.A. (2014). Diversity and biotechnological potential of microorganisms associated with marine sponges. Appl. Microbiol. Biotechnol. *98*, 7331–7347.

Fukunaga, Y., Kurahashi, M., Tanaka, K., Yanagi, K., Yokota, A., and Harayama, S. (2006). Pseudovibrio ascidiaceicola sp. nov., isolated from ascidians (sea squirts). Int. J. Syst. Evol. Microbiol. *56*, 343–347.

Fuqua, W.C., Winans, S.C., and Greenberg, E.P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. J. Bacteriol. *176*, 269–275.

Gazave, E., Lapébie, P., Renard, E., Vacelet, J., Rocher, C., Ereskovsky, A.V., Lavrov, D.V., and Borchiellini, C. (2010). Molecular phylogeny restores the supra-generic subdivision of homoscleromorph sponges (Porifera, Homoscleromorpha). PloS One *5*, e14290.

Gazave, E., Lapébie, P., Ereskovsky, A.V., Vacelet, J., Renard, E., Cárdenas, P., and Borchiellini, C. (2012). No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. Hydrobiologia *687*, 3–10.

Guy, L., Nystedt, B., Toft, C., Zaremba-Niedzwiedzka, K., Berglund, E.C., Granberg, F., Näslund, K., Eriksson, A.-S., and Andersson, S.G.E. (2013). A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen Bartonella. PLoS Genet. *9*, e1003393.

Habyarimana, F., Al-Khodor, S., Kalia, A., Graham, J.E., Price, C.T., Garcia, M.T., and Kwaik, Y.A. (2008). Role for the Ankyrin eukaryotic-like genes of Legionella pneumophila in parasitism of protozoan hosts and human macrophages. Environ. Microbiol. *10*, 1460–1474.

Hallam, S.J., Konstantinidis, K.T., Putnam, N., Schleper, C., Watanabe, Y., Sugahara, J., Preston, C., Torre, J. de la, Richardson, P.M., and DeLong, E.F. (2006). Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum. Proc. Natl. Acad. Sci. *103*, 18296–18301.

Harjes, J., Ryu, T., Abdelmohsen, U.R., Moitinho-Silva, L., Horn, H., Ravasi, T., and Hentschel, U. (2014). Draft Genome Sequence of the Antitrypanosomally Active Sponge-Associated Bacterium Actinokineospora sp. Strain EG49. Genome Announc. *2*, e00160–14.

Heise, T., and Dersch, P. (2006). Identification of a domain in Yersinia virulence factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake. Proc. Natl. Acad. Sci. U. S. A. *103*, 3375–3380.

Hentschel, U., Schmid, M., Wagner, M., Fieseler, L., Gernert, C., and Hacker, J. (2001). Isolation and phylogenetic analysis of bacteria with antimicrobial activities from the Mediterranean sponges Aplysina aerophoba and Aplysina cavernicola. FEMS Microbiol. Ecol. *35*, 305–312.

Hentschel, U., Piel, J., Degnan, S.M., and Taylor, M.W. (2012). Genomic insights into the marine sponge microbiome. Nat. Rev. Microbiol. *10*, 641–654.

Hill, M., and Wilcox, T. (1998). Unusual mode of symbiont repopulation after bleaching in Anthosigmella varians: acquisition of different zooxanthellae strains. Symbiosis Phila. PaUSA.

Hirose, E., and Fukuda, T. (2006). Vertical transmission of photosymbionts in the colonial ascidian Didemnum molle: the larval tunic prevents symbionts from attaching to the anterior part of larvae. Zoolog. Sci. *23*, 669–674.

Hoffmann, F., Larsen, O., Thiel, V., Rapp, H.T., Pape, T., Michaelis, W., and Reitner, J. (2005). An anaerobic world in sponges. Geomicrobiol. J. *22*, 1–10.

Hosoya, S., and Yokota, A. (2007). Pseudovibrio japonicus sp. nov., isolated from coastal seawater in Japan. Int. J. Syst. Evol. Microbiol. *57*, 1952–1955.

Hu, G.-P., Yuan, J., Sun, L., She, Z.-G., Wu, J.-H., Lan, X.-J., Zhu, X., Lin, Y.-C., and Chen, S.-P. (2011). Statistical research on marine natural products based on data obtained between 1985 and 2008. Mar. Drugs *9*, 514–525.

Imhoff, J.F., and Trüper, H.G. (1976). Marine sponges as habitats of anaerobic phototrophic bacteria. Microb. Ecol. *3*, 1–9.

Iturbe-Ormaetxe, I., Burke, G.R., Riegler, M., and O'Neill, S.L. (2005). Distribution, expression, and motif variability of ankyrin domain genes in Wolbachia pipientis. J. Bacteriol. *187*, 5136–5145.

Jani, A.J., and Cotter, P.A. (2010). Type VI Secretion: Not Just for Pathogenesis Anymore. Cell Host Microbe *8*, 2–6.

Joullié, M.M., Leonard, M.S., Portonovo, P., Liang, B., Ding, X., and La Clair, J.J. (2003). Chemical Defense in Ascidians of the Didemnidae Family†. Bioconjug. Chem. *14*, 30–37.

Juhas, M. (2015). Type IV secretion systems and genomic islands-mediated horizontal gene transfer in Pseudomonas and Haemophilus. Microbiol. Res. *170*, 10–17.

Kapitein, N., and Mogk, A. (2013). Deadly syringes: type VI secretion system activities in pathogenicity and interbacterial competition. Curr. Opin. Microbiol. *16*, 52–58.

Kennedy, J., Baker, P., Piper, C., Cotter, P.D., Walsh, M., Mooij, M.J., Bourke, M.B., Rea, M.C., O'Connor, P.M., Ross, R.P., et al. (2009). Isolation and analysis of bacteria with antimicrobial activities from the marine sponge Haliclona simulans collected from Irish waters. Mar. Biotechnol. N. Y. N *11*, 384–396.

Kersters, K., and Ley, J.D. (1968). The occurrence of the Entner-Doudoroff pathway in bacteria. Antonie Van Leeuwenhoek *34*, 393–408.

Kikuchi, T., Mizunoe, Y., Takade, A., Naito, S., and Yoshida, S. (2005). Curli fibers are required for development of biofilm architecture in Escherichia coli K-12 and enhance bacterial adherence to human uroepithelial cells. Microbiol. Immunol. *49*, 875–884.

Knöbl, T., Baccaro, M.R., Moreno, A.M., Gomes, T.A., Vieira, M.A., Ferreira, C.S., and Ferreira, A.J. (2001). Virulence properties of Escherichia coli isolated from ostriches with respiratory disease. Vet. Microbiol. *83*, 71–80.

Koopmans, M., and Wijffels, R.H. (2008). Seasonal growth rate of the sponge Haliclona oculata (Demospongiae: Haplosclerida). Mar. Biotechnol. N. Y. N *10*, 502–510.

Koren, O., and Rosenberg, E. (2006). Bacteria Associated with Mucus and Tissues of the Coral Oculina patagonica in Summer and Winter. Appl. Environ. Microbiol. *72*, 5254–5259.

Kumar, S., Subramanian, S., Raghava, G.P.S., and Pinnaka, A.K. (2012). Genome Sequence of the Marine Bacterium Marinilabilia salmonicolor JCM 21150T. J. Bacteriol. *194*, 3746–3746.

Kwan, J.C., and Schmidt, E.W. (2013). Bacterial endosymbiosis in a chordate host: long-term co-evolution and conservation of secondary metabolism. PloS One *8*, e80822.

Kwan, J.C., Donia, M.S., Han, A.W., Hirose, E., Haygood, M.G., and Schmidt, E.W. (2012). Genome streamlining and chemical defense in a coral reef symbiosis. Proc. Natl. Acad. Sci. U. S. A. *109*, 20655–20660.

Lafi, F.F., Fuerst, J.A., Fieseler, L., Engels, C., Goh, W.W.L., and Hentschel, U. (2009). Widespread Distribution of Poribacteria in Demospongiae. Appl Env. Microbiol *75*, 5695–5699.

Lafont, J.P., Dho, M., D'Hauteville, H.M., Bree, A., and Sansonetti, P.J. (1987). Presence and expression of aerobactin genes in virulent avian strains of Escherichia coli. Infect. Immun. *55*, 193–197.

Laport, M.S., Santos, O.C.S., and Muricy, G. (2009). Marine sponges: potential sources of new antimicrobial drugs. Curr. Pharm. Biotechnol. *10*, 86–105.

Lee, O.O., Wang, Y., Yang, J., Lafi, F.F., Al-Suwailem, A., and Qian, P.-Y. (2011). Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. ISME J. *5*, 650–664.

Leys, S.P., Mackie, G.O., and Reiswig, H.M. (2007). The biology of glass sponges. Adv. Mar. Biol. *52*, 1–145.

Ling, J., Pan, H., Gao, Q., Xiong, L., Zhou, Y., Zhang, D., Gao, S., and Liu, X. (2013). Aerobactin Synthesis Genes iucA and iucC Contribute to the Pathogenicity of Avian Pathogenic Escherichia coli O2 Strain E058. PLoS ONE *8*.

Liu, M., Fan, L., Zhong, L., Kjelleberg, S., and Thomas, T. (2012). Metaproteogenomic analysis of a community of sponge symbionts. ISME J. *6*, 1515–1525.

Liu, M.Y., Kjelleberg, S., and Thomas, T. (2011). Functional genomic analysis of an uncultured ?-proteobacterium in the sponge Cymbastela concentrica. ISME J. *5*, 427–435.

Lührmann, A., Nogueira, C.V., Carey, K.L., and Roy, C.R. (2010). Inhibition of pathogen-induced apoptosis by a Coxiella burnetii type IV effector protein. Proc. Natl. Acad. Sci. U. S. A. *107*, 18997–19001.

Mahon, A.R., Amsler, C.D., McClintock, J.B., Amsler, M.O., and Baker, B.J. (2003). Tissue-specific palatability and chemical defenses against macropredators and pathogens in the common articulate brachiopod Liothyrella uva from the Antarctic Peninsula. J. Exp. Mar. Biol. Ecol. *290*, 197–210.

Mavromatis, K., Doyle, C.K., Lykidis, A., Ivanova, N., Francino, M.P., Chain, P., Shin, M., Malfatti, S., Larimer, F., Copeland, A., et al. (2006). The genome of the obligately intracellular bacterium Ehrlichia canis reveals themes of complex membrane structure and immune evasion strategies. J. Bacteriol. *188*, 4015–4023.

Mittl, P.R.E., and Schneider-Brachert, W. (2007). Sel1-like repeat proteins in signal transduction. Cell. Signal. *19*, 20–31.

Molinski, T.F., Dalisay, D.S., Lievens, S.L., and Saludes, J.P. (2009). Drug development from marine natural products. Nat. Rev. Drug Discov. *8*, 69–85.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordoñez, C.L., Lory, S., et al. (2006). A virulence locus of

Pseudomonas aeruginosa encodes a protein secretion apparatus. Science *312*, 1526–1530.

Mukherjee, S., Keitany, G., Li, Y., Wang, Y., Ball, H.L., Goldsmith, E.J., and Orth, K. (2006). Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. Science *312*, 1211–1214.

Muller-Parker, G., and D'Elia, C.F. (1997). Interactions between corals and their symbiotic algae. Life Death Coral Reefs 96–113.

Neufeld, J.D., Wagner, M., and Murrell, J.C. (2007). Who eats what, where and when? Isotope-labelling experiments are coming of age. ISME J. *1*, 103–110.

Ng, W.-L., and Bassler, B.L. (2009). Bacterial quorum-sensing network architectures. Annu. Rev. Genet. *43*, 197–222.

Niebuhr, K., Jouihri, N., Allaoui, A., Gounon, P., Sansonetti, P.J., and Parsot, C. (2000). IpgD, a protein secreted by the type III secretion machinery of Shigella flexneri, is chaperoned by IpgE and implicated in entry focus formation. Mol. Microbiol. *38*, 8–19.

Nougayrède, J.-P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G., Buchrieser, C., Hacker, J., Dobrindt, U., and Oswald, E. (2006). Escherichia coli induces DNA double-strand breaks in eukaryotic cells. Science *313*, 848–851.

O'Halloran, J.A., Barbosa, T.M., Morrissey, J.P., Kennedy, J., O'Gara, F., and Dobson, A.D.W. (2011). Diversity and antimicrobial activity of Pseudovibrio spp. from Irish marine sponges. J. Appl. Microbiol. *110*, 1495–1508.

Olsén, A., Jonsson, A., and Normark, S. (1989). Fibronectin binding mediated by a novel class of surface organelles on Escherichia coli. Nature *338*, 652–655.

Olsén, A., Herwald, H., Wikström, M., Persson, K., Mattsson, E., and Björck, L. (2002). Identification of Two Protein-binding and Functional Regions of Curli, a Surface Organelle and Virulence Determinant ofEscherichia coli. J. Biol. Chem. *277*, 34568–34572.

Pan, X., Lührmann, A., Satoh, A., Laskowski-Arce, M.A., and Roy, C.R. (2008). Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. Science *320*, 1651–1654.

Patankar, A.V., and González, J.E. (2009). Orphan LuxR regulators of quorum sensing. FEMS Microbiol. Rev. *33*, 739–756.

Patra, T., Koley, H., Ramamurthy, T., Ghose, A.C., and Nandy, R.K. (2012). The Entner-Doudoroff Pathway Is Obligatory for Gluconate Utilization and Contributes to the Pathogenicity of Vibrio cholerae. J. Bacteriol. *194*, 3377–3385.

Paul, V.J., and Puglisi, M.P. (2004). Chemical mediation of interactions among marine organisms. Nat. Prod. Rep. *21*, 189–209.

Paul, V.J., Lindquist, N., and Fenical, W. (1990). Chemical defenses of the tropical ascidian Atapozoa sp. and its nudibranch predators Nembrotha spp. Mar. Ecol. Prog. Ser. Oldendorf *59*, 109–118.

Paul, V.J., Cruz-Rivera, E., and Thacker, R.W. (2001). Chemical mediation of macroalgal-herbivore interactions: ecological and evolutionary perspectives. Mar. Chem. Ecol. CRC Press Boca Raton 227–265.

Paul, V.J., Puglisi, M.P., and Ritson-Williams, R. (2006). Marine chemical ecology. Nat. Prod. Rep. *23*, 153–180.

Paul, V.J., Ritson-Williams, R., and Sharp, K. (2011). Marine chemical ecology in benthic environments. Nat. Prod. Rep. *28*, 345–387.

Penesyan, A., Tebben, J., Lee, M., Thomas, T., Kjelleberg, S., Harder, T., and Egan, S. (2011). Identification of the antibacterial compound produced by the marine epiphytic bacterium Pseudovibrio sp. D323 and related sponge-associated bacteria. Mar. Drugs *9*, 1391–1402.

Planet, P.J., Kachlany, S.C., Fine, D.H., DeSalle, R., and Figurski, D.H. (2003). The Widespread Colonization Island of Actinobacillus actinomycetemcomitans. Nat. Genet. *34*, 193–198.

Preston, G.M. (2007). Metropolitan microbes: type III secretion in multihost symbionts. Cell Host Microbe *2*, 291–294.

Proksch, P. (1994). Defensive roles for secondary metabolites from marine sponges and sponge-feeding nudibranchs. Toxicon *32*, 639–655.

Rath, C.M., Janto, B., Earl, J., Ahmed, A., Hu, F.Z., Hiller, L., Dahlgren, M., Kreft, R., Yu, F., Wolff, J.J., et al. (2011). Meta-omic characterization of the marine invertebrate

microbial consortium that produces the chemotherapeutic natural product ET-743. ACS Chem. Biol. *6*, 1244–1256.

Riesenfeld, C.S., Murray, A.E., and Baker, B.J. (2008). Characterization of the microbial community and polyketide biosynthetic potential in the palmerolide-producing tunicate Synoicum adareanum. J. Nat. Prod. *71*, 1812–1818.

Rolain, J.-M., Foucault, C., Guieu, R., La Scola, B., Brouqui, P., and Raoult, D. (2002). Bartonella quintana in human erythrocytes. Lancet *360*, 226–228.

Sagar, S., Kaur, M., and Minneman, K.P. (2010). Antiviral lead compounds from marine sponges. Mar. Drugs *8*, 2619–2638.

Santos, O.C.S., Pontes, P.V.M.L., Santos, J.F.M., Muricy, G., Giambiagi-deMarval, M., and Laport, M.S. (2010). Isolation, characterization and phylogeny of sponge-associated bacteria with antimicrobial activities from Brazil. Res. Microbiol. *161*, 604–612.

Schmidt, E.W., and Donia, M.S. (2009). Cyanobactin ribosomally synthesized peptides — a case of deep metagenome mining. Methods Enzymol. *458*, 575–596.

Schmitt, S., Hentschel, U., and Taylor, M. (2012). Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges. Hydrobiologia *687*, 341–351.

Schmitz-Esser, S., Tischler, P., Arnold, R., Montanaro, J., Wagner, M., Rattei, T., and Horn, M. (2010). The Genome of the Amoeba Symbiont “Candidatus Amoebophilus asiaticus” Reveals Common Mechanisms for Host Cell Interaction among Amoeba-Associated Bacteria. J. Bacteriol. *192*, 1045–1057.

Schreiner, H.C., Sinatra, K., Kaplan, J.B., Furgang, D., Kachlany, S.C., Planet, P.J., Perez, B.A., Figurski, D.H., and Fine, D.H. (2003). Tight-adherence genes of Actinobacillus actinomycetemcomitans are required for virulence in a rat model. Proc. Natl. Acad. Sci. U. S. A. *100*, 7295–7300.

Schwarzer, D., Finking, R., and Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. Nat. Prod. Rep. *20*, 275–287.

Sertan-de Guzman, A.A., Predicala, R.Z., Bernardo, E.B., Neilan, B.A., Elardo, S.P., Mangalindan, G.C., Tasdemir, D., Ireland, C.M., Barraquio, W.L., and Concepcion,

G.P. (2007). Pseudovibrio denitrificans strain Z143-1, a heptylprodigiosin-producing bacterium isolated from a Philippine tunicate. FEMS Microbiol. Lett. *277*, 188–196.

Sheng, L., Lv, Y., Liu, Q., Wang, Q., and Zhang, Y. (2013). Connecting type VI secretion, quorum sensing, and c-di-GMP production in fish pathogen Vibrio alginolyticus through phosphatase PppA. Vet. Microbiol. *162*, 652–662.

Shenkar, N., and Swalla, B.J. (2011). Global Diversity of Ascidiacea. PLoS ONE *6*, e20657.

Siegl, A., Kamke, J., Hochmuth, T., Piel, J., Richter, M., Liang, C., Dandekar, T., and Hentschel, U. (2011). Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. ISME J. *5*, 61–70.

Silva, V. (2012). Proto-Eukaryote Symbiosis. Retrieved from University of Porto Repository. Bachelor report in Biochemistry

Staunton, J., and Weissman, K.J. (2001). Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. *18*, 380–416.

Steger, D., Ettinger-Epstein, P., Whalan, S., Hentschel, U., De Nys, R., Wagner, M., and Taylor, M.W. (2008). Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges. Environ. Microbiol. *10*, 1087–1094.

Tang, K.-H., Feng, X., Tang, Y.J., and Blankenship, R.E. (2009). Carbohydrate Metabolism and Carbon Fixation in Roseobacter denitrificans OCh114. PLoS ONE *4*, e7233.

Taylor, M.W., Radax, R., Steger, D., and Wagner, M. (2007). Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. Microbiol. Mol. Biol. Rev. *71*, 295–347.

Thacker, R.W. (2005). Impacts of Shading on Sponge-Cyanobacteria Symbioses: A Comparison between Host-Specific and Generalist Associations. Integr. Comp. Biol. *45*, 369–376.

Thakur, N.L., Hentschel, U., Krasko, A., Pabel, C.T., Anil, A.C., and Mller, W.E.G. (2003). Antibacterial activity of the sponge Suberites domuncula and its primmorphs: potential basis for epibacterial chemical defense. Aquat. Microb. Ecol. *31*, 77–83.

Thiel, V., and Imhoff, J.F. (2003). Phylogenetic identification of bacteria with antimicrobial activities isolated from Mediterranean sponges. Biomol. Eng. *20*, 421–423.

Thomas, T., Rusch, D., DeMaere, M.Z., Yung, P.Y., Lewis, M., Halpern, A., Heidelberg, K.B., Egan, S., Steinberg, P.D., and Kjelleberg, S. (2010). Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. ISME J. *4*, 1557–1567.

Thoms, C., Horn, M., Wagner, M., Hentschel, U., and Proksch, P. (2003). Monitoring microbial diversity and natural product profiles of the sponge Aplysina cavernicola following transplantation. Mar. Biol. *142*, 685–692.

Tomich, M., Planet, P.J., and Figurski, D.H. (2007). The tad locus: postcards from the widespread colonization island. Nat. Rev. Microbiol. *5*, 363–375.

Unson, M.D., Holland, N.D., and Faulkner, D.J. (1994). A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. Mar. Biol. *119*, 1–11.

Van Soest, R.W.M., Boury-Esnault, N., Vacelet, J., Dohrmann, M., Erpenbeck, D., De Voogd, N.J., Santodomingo, N., Vanhoorne, B., Kelly, M., and Hooper, J.N.A. (2012). Global Diversity of Sponges (Porifera). PLoS ONE *7*.

Vogel, G. (2008). The inner lives of sponges. Science *320*, 1028–1030.

Voth, D.E., Howe, D., Beare, P.A., Vogel, J.P., Unsworth, N., Samuel, J.E., and Heinzen, R.A. (2009). The Coxiella burnetii Ankyrin Repeat Domain-Containing Protein Family Is Heterogeneous, with C-Terminal Truncations That Influence Dot/Icm-Mediated Secretion. J. Bacteriol. *191*, 4232–4242.

Voth, D.E., Broederdorf, L.J., and Graham, J.G. (2012). Bacterial Type IV secretion systems: versatile virulence machines. Future Microbiol. *7*, 241–257.

Waldor, M.K. (2010). Mobilizable genomic islands: going mobile with oriT mimicry. Mol. Microbiol. *78*, 537–540.

Webster, N.S., Luter, H.M., Soo, R.M., Botté, E.S., Simister, R.L., Abdo, D., and Whalan, S. (2012). Same, same but different: symbiotic bacterial associations in GBR sponges. Front. Microbiol. *3*, 444.

Wiley, D.J., Nordfeldth, R., Rosenzweig, J., DaFonseca, C.J., Gustin, R., Wolf-Watz, H., and Schesser, K. (2006). The Ser/Thr kinase activity of the Yersinia protein kinase A (YpkA) is necessary for full virulence in the mouse, mollifying phagocytes, and disrupting the eukaryotic cytoskeleton. Microb. Pathog. *40*, 234–243.

Wiley, D.J., Shrestha, N., Yang, J., Atis, N., Dayton, K., and Schesser, K. (2009). The Activities of the Yersinia Protein Kinase A (YpkA) and Outer Protein J (YopJ) Virulence Factors Converge on an eIF2α Kinase. J. Biol. Chem. *284*, 24744–24753.

Wilkinson, C.R., and Fay, P. (1979). Nitrogen fixation in coral reef sponges with symbiotic cyanobacteria. Nature *279*, 527–529.

Zhaxybayeva, O., and Doolittle, W.F. (2011). Lateral gene transfer. Curr. Biol. *21*, R242–R246.

Guide to the Third Edition of the OED.